Published in final edited form as:
Chemistry. 2009 ; 15(15): 3761–3772. doi:10.1002/chem.200802105.

The Correlation of ^{113}Cd NMR and ^{111m}Cd PAC Spectroscopies Provides a Powerful Approach for the Characterization of the Structure of Cd^{II}-Substituted Zn^{II} Proteins**

Olga Iranzo[a], **Tamas Jakusch**[a], **Kyung-Hoon Lee**[a], **Lars Hemmingsen**[b], and **Vincent L. Pecoraro**[*,a]

[a]Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109-1055 (USA)

[b]Bioinorganic Chemistry Group, IGM, Faculty of Life Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg C (Denmark)

Abstract

Cd^{II} has been used as a probe of zinc metalloenzymes and proteins because of the spectroscopic silence of Zn^{II}. One of the most commonly used spectroscopic techniques is ^{113}Cd NMR; however, in recent years ^{111m}Cd Perturbed Angular Correlation spectroscopy (^{111m}Cd PAC) has also been shown to provide useful structural, speciation and dynamics information on Cd^{II} complexes and biomolecules. In this article, we show how the joint use of ^{113}Cd NMR and ^{111m}Cd PAC spectroscopies can provide detailed information about the Cd^{II} environment in thiolate-rich proteins. Specifically we show that the ^{113}Cd NMR chemical shifts observed for Cd^{II} in the designed **TRI** series (TRI = Ac-G-(LKALEEK)$_4$G-NH_2) of peptides vary depending on the proportion of trigonal planar CdS_3 and pseudotetrahedral CdS_3O species present in the equilibrium mixture. PAC spectra are able to quantify these mixtures. When one compares the chemical shift range for these peptides (from δ = 570 to 700 ppm), it is observed that CdS_3 species have δ 675–700 ppm, CdS_3O complexes fall in the range δ 570–600 ppm and mixtures of these forms fall linearly between these extremes. If one then determines the pK_{a2} values for Cd^{II} complexation [pK_{a2} is for the reaction $Cd[(peptide–H)_2(peptide)]^+ \rightarrow Cd\text{-}(peptide)_3^- + 2H^+$ and compares these to the observed chemical shift for the $Cd(peptide)_3^-$ complexes, one finds that there is also a direct linear correlation. Thus, by determining the chemical shift value of these species, one can directly assess the metal-binding affinity of the construct. This illustrates how proteins may be able to fine tune metal-binding affinity by destabilizing one metallospecies with respect to another. More important, these studies demonstrate that one may have a broad ^{113}Cd NMR chemical shift range for a chemical species (e.g., CdS_3O) which is not necessarily a reflection of the structural diversity within such a four-coordinate species, but rather a consequence of a fast exchange equilibrium between two related species (e.g., CdS_3O and CdS_3). This could lead to reinterpretation of the assignments of cadmium–protein complexes and may impact the application of Cd^{II} as a probe of Zn^{II} sites in biology.

Keywords

cadmium; metallopeptides; NMR spectroscopy; zinc proteins

PAC Spectroscopy = perturbed angular correlation spectroscopy Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/chem.200802105: Models and equations for extended model calculations in order to fit the UV/Vis pH titration curves; pK for several equilibria determined for several peptides: 1) **TRIL16Pen, 3) **GRANDL16Pen**, 4) **TRIL16CL23A**, 6) **BABYL9C**, 10) **GRANDL9C**, 12) **TRIL9AL16C**, 13) **TRIL12AL16Pen**, 14) **TRIL12AL16C**.

Fax: (+1) 734-936-7628, vlpec@umich.edu.

Introduction

Zinc proteins are among the most abundant metalloproteins, with the metal center either having a catalytic role or a structural role. These ubiquitous metalloproteins are involved in extraordinarily diverse processes which extend from hydrolytic reactions in metabolism to gene regulation (transcription factors), DNA repair and post-transcriptional modification of proteins.[1–10] In addition, the advent of proteomics is increasing our recognition of the number of proteins that require Zn^{II} either as a catalytic and/or structural element.[11] This utilization is not surprising since zinc is an abundant trace metal, found second only to iron in eukaryotic organisms. Therefore, to clarify fully metallobiochemistry, one needs a deep understanding of the structural aspects defining different zinc-binding sites and how the protein framework proficiently modulates the properties of this metal ion to perform, in each case, the required function.

Unlike other common metals in biology, zinc is promiscuous with respect to the types and number of ligands bound to it in proteins. One can observe structures that are low coordination number (e.g., 4), sulfur-rich centers to moderate coordination number (5 or 6) environments that are oxygen and nitrogen rich. Unfortunately, due to the poor spectroscopic properties of Zn^{II}, a diamagnetic closed shell d^{10} ion, it is difficult to obtain information regarding the structure of these binding sites and, in the case of enzymes, of their mechanisms of action. To overcome this drawback, scientists have widely replaced Zn^{II} with other metal ions that have useful spectroscopic properties.[12,13]

Cd^{II} has often been employed as a probe to understand Zn^{II}-binding sites. This is because Cd^{II}-coordination preferences resemble the broad coordination number and ligand affinities of Zn^{II}, but permits the use of ^{113}Cd NMR spectroscopy—an extremely useful and accessible technique that is very sensitive to the type and number of ligands bound to Cd^{II}.[12,14,15] Less used, but gaining recognition, is ^{111m}Cd Perturbed Angular Correlation (PAC) spectroscopy.[16] Until recently, ^{113}Cd NMR and ^{111m}Cd PAC spectroscopies were rarely used in combination to study natural metal-binding sites, and in fact, few examples exist in the literature.[17–19] Both methods are highly sensitive to the Cd^{II} first coordination sphere and provide complementary dynamics information because of the different timescale of the methods. Thus, the combined use of these two techniques could be a valuable strategy to probe Cd^{II}-substituted Zn^{II}-binding sites in biomolecules; however, the expectation for establishing a proper structural interpretation can only be realized once a systematic correlation of the ^{113}Cd NMR and ^{111m}Cd PAC signatures of Cd^{II} in well defined environments exists.

Our group has used a de novo design approach to define the chemistry of Cd^{II} in a thiolate-rich environment as might be found in CadC[20] or cadmium substituted Zn^{II} proteins, such as aminolevulinic acid dehydratase (ALAD)[21] or the hepatitis C virus NS_3 protease domain.[22] The **TRI** peptide family (TRI = Ac-G(LKALEEK)$_4$G-NH_2, Table 1)[23,24] utilizes a heptad repeat unit that contains hydrophobic residues, usually Leu, in the 1st (“**a**”) and 4th (“**d**”) positions, generating an amphiphatic a-helix that in aqueous solution associates above pH 6.0 to form three-stranded coiled coils. We have been studying the binding of Cd^{II} to different **TRI** peptides where a single Leu, either in position “**a**” or “**d**”, was replaced by a Cys to generate a thiol-rich binding site inside the hydrophobic core of these coiled coils.[25–29] These peptides are the perfect scaffolds as they are much simpler than the natural systems but retain enough structural complexity to study how the protein framework can modulate the geometries and physical properties of metal ions bound to thiol-rich centers. By combining ^{113}Cd NMR and ^{111m}Cd PAC spectroscopies, we have shown how the peptide **TRIL16C** binds Cd^{II} as a mixture of pseudotetrahedral (CdS_3O) and

trigonal planar (CdS_3) structures.[25] Subtle changes in the amino acid sequence of the parent peptide were sufficient to generate two peptides, **TRIL12AL16C** and **TRIL16Pen**, that bind Cd^{II} exclusively in the CdS_3O and CdS_3 geometries, respectively.[30] Furthermore, these techniques have driven our design strategy for the heterochromic peptide, **GRANDL16PenL26AL30C**, a single peptide containing two binding sites within the same three-stranded coiled coil which is capable of binding two Cd^{II} with different physical properties and showing site-selective Cd^{II} recognition.[31] The combination of the ^{113}Cd NMR and ^{111m}Cd PAC spectroscopic data has been crucial to the characterization of these peptides to high chemical resolution.

The ^{113}Cd NMR and ^{111m}Cd PAC spectroscopic data also provide useful information on the chemical exchange because of the different time scales probed by these techniques [NMR (0.01–10 ms); ^{111m}Cd PAC (0.1–100 ns).[25,28] For a peptide, such as **TRIL16C**, a single resonance (625 ppm) is observed suggesting that a single Cd structure is present; however, PAC studies revealed the existence of the CdS_3O and CdS_3 species. Based on these observations, we have been able to place a limitation on the exchange rate for the oxygen of the CdS_3O species as being between 0.01–10 ms. More important, these data reveal an important caveat when using ^{113}Cd NMR as the sole method of defining the metal-coordination sphere. In cases, where ligand exchange occurs, the chemical shift range for the ^{113}Cd NMR may not alone be representative of the true speciation of a system.

Here we present a systematic study of CdS_3O and CdS_3 centers in selected peptides that are analyzed by ^{113}Cd NMR, ^{111m}Cd PAC and UV/Vis spectroscopies. Important linear correlations are observed between ^{113}Cd NMR and ^{111m}Cd PAC spectroscopies and the acid/base properties of the metal-binding site that allows one to identify the proportion of three- and four-coordinate sites in any peptide. These types of correlations are very hard to obtain in native systems since they require a system capable of accommodating different metal coordination geometries while keeping basically the same overall structure. Indeed, to the best of our knowledge, there are no reports in the literature where this type of data, ^{111m}Cd PAC–^{113}Cd NMR–Cd^{II} binding acidity, has been correlated either for native systems or small molecules. Since numerous zinc proteins bind Zn^{II} with thiolate ligands,[7,9,21,22,32–37] and Cd^{II} substitution is quite often used to study these binding sites, the results reported here can shed light into understanding better Cd^{II}-substituted thiolate-rich zinc proteins.

Results

^{111m}Cd Perturbed Angular Correlation (PAC) spectroscopy

^{111m}Cd PAC spectroscopy was used to determine the Cd^{II} coordination geometry of the different **TRI** peptides complexes. The ^{111m}Cd PAC spectra for different peptides of the **TRI** family are shown in Figure 1. The PAC data were analyzed as described in the Experimental Section. Each nuclear quadrupole interaction (NQI) was modelled using a separate set of parameters that includes ω_0, η, $\Delta\omega_0/\omega_0$, $1/\tau_c$ and A. The parameters fitted to the PAC data for each peptide complex are reported in Table 2. All the NQIs have relatively low values of η, the asymmetry parameter, indicating that there is symmetry around the z axis which is the principal axis of the electric field gradient tensor with the largest eigenvalue. All the values of ω_0, the parameter giving information regarding the first coordination sphere ligands, fall into two frequency regions, one around 0.450 rad ns^{-1} and the other around 0.340 rad ns^{-1}. This indicates that two coordination geometries, trigonal planar (ω_0 = 0.450 rad ns^{-1}) and distorted tetrahedral (ω_0 = 0.340 rad ns^{-1}), may adequately describe the metal site structures in all these peptides. The PAC spectroscopic data for the peptides **TRIL9AL16C** and **TRIL12AL16Pen** give in both cases a well defined NQI with ω_0 and η values of 0.3332 rad ns^{-1} and 0.19, and 0.339 rad ns^{-1} and 0.36, respectively.

These values are similar to the value observed for the peptide **TRIL12AL16C** with a ω_0 = 0.3405 rad ns^{-1} and η = 0.141,[30] and may correspond to a distorted tetrahedral CdS_3O species. The NQI for the peptide containing the Pen substitution, **TRIL12AL16Pen**, has the highest value of η (0.36) in comparison with the other two peptides which is indicative of a larger perturbation of the geometry from axial symmetry. The peptides **TRIL16-CL23A** and **TRIL12VL16Pen** give PAC spectra showing two NQIs which are similar to those found for the peptide **TRIL16C**.[25] These results indicate that Cd^{II} binds to these peptides as a mixture of distorted tetrahedral CdS_3O species (lower frequency species) and distorted trigonal planar CdS_3 species (higher frequency species). For the peptide **TRIL16-CL23A**, there is a similar occupation of the two geometries since the amplitude of both frequencies is similar. However, for **TRIL12VL16Pen**, there is a dominating species and this belongs to the high-frequency group. The low-frequency component is difficult to fit due to the low amplitude of this signal. Thus, in order to limit the number of free parameters fitted to this NQI, the ω_0, η, $\Delta\omega_0/\omega_0$ and $1/\tau_c$ values were fixed to the values of low frequency component found for the **TRIL16C** peptide. Different parameters fitted to the low frequency signal may give equally good χ_r^2 values. A relative high $1/\tau_c$ value is observed for the signals in all the peptides containing the Pen substitution species (Table 2, see last three rows). This is indicative of local dynamics at the metal ion binding site that occur on the nanosecond timescale. For **TRIL16CL23A** an experiment was carried out without TRIS buffer and no significant change of the spectrum was observed.

UV/Vis spectroscopy

The pH dependence of Cd^{II} binding to the thiol groups of different peptides was followed by monitoring the formation of the characteristic ligand to metal charge transfer (LMCT) band at 235 nm.[27,29] For those experiments where Cd^{II} was binding to the peptides **TRIL16Pen, GRANDL16Pen** and **TRIL12VL16Pen**, longer equilibration times after addition of KOH were in general required. Figure 2 shows the pH profiles for several of the peptides of the **TRI** family. The experimental data was fit using the model published previously that corresponds to the simultaneous release of the two protons (see Experimental Section).[27] This model implies the formation of the species $Cd[(peptide–H)_2(peptide)]^+$ at low pH. The apparent acidity constant K_{a2} has thus units of molar squared (M^2) and the apparent pK_{a2} values determined for each peptide are reported in Table 3. The experimental data was also fit using an extended model that did not imply the full formation of the species $Cd[(peptide–H)_2(peptide)]^+$ and included different equilibria (see Supporting Information). Although the values obtained for pK_{a2} are slightly different, the information they provide is identical and the same trend is observed (see Discussion and Figure S1).

^{113}Cd NMR spectroscopy

The ^{113}Cd NMR spectra of the different Cd^{II} peptide complexes were recorded at pH values where, based on the pH titrations, the metal ion was fully bound (pH 8.5–9.0). In all cases, a single resonance is observed and the ^{113}Cd chemical shifts obtained (Table 3) are consistent with thiol coordination to Cd^{II}.[14,38]

Discussion

Proteins harness the chemical properties of Zn^{II} in order to perform a large variety of important cellular functions. The most common zinc binding sites are distorted tetrahedral Zn^{II} centers where the metal ion is typically bound to nitrogen, oxygen and sulfur donor ligands. Usually the amino acids defining these binding pockets are His (N), Cys (S), Glu (O) and Asp (O), with the first two amino acids being the most prevalent donors.[1,39] Thiolate-rich binding centers are found in proteins implicated in very different biological activities, including gene regulation, methyl transfer reactions and protein

modification.[7,9,10,32–34,40,41] Surprisingly, despite possessing a common tetrahedral geometry and similar binding pockets, these thiolaterich Zn^{II} centers are accurately tuned to carry out their distinct biological function. Therefore, it has become clear that besides the chemical nature of the coordinating amino acids, there are additional features of the metal environment which control the properties of Zn^{II}. Unfortunately, a clear knowledge of the structural, thermodynamic and kinetic factors that differentially modulate the chemical and physical properties of Zn^{II} bound to these sites is sparse. This lack of information stems, to some extent, from the fact that Zn^{II} is not amenable to most forms of spectroscopies, which complicates the direct study of the zinc-binding properties of these proteins. The use of Cd^{II} as a probe to study naturally containing Zn^{II} proteins has been one of the most employed alternatives, since this metal replacement opens the door to the use of ^{113}Cd NMR and ^{111m}Cd PAC spectroscopies.[12,14,16]

Our group is now in the position to start analyzing in great detail specific thiol-rich binding sites using de novo designed peptides. These peptides sequester Cd^{II} by enforcing a pseudotetrahedral S_3O first coordination sphere. The fourth ligand, thought to be an exogenous water molecule, can be present or not, generating in the last case a trigonal planar S_3 coordination site. The significance of studying these systems lays in the fact that in zinc enzymes, the Zn^{II} center normally possesses three permanent coordinating positions occupied by amino acid ligands and the fourth site is filled either with an exogenous water/hydroxide molecule, which is very important for catalysis, a fourth amino acid or serves as the location of substrate binding.[1–3,22,34,35,42]

We have successfully designed two peptides, **TRIL12AL16C** and **TRIL16Pen**, capable of binding Cd^{II} exclusively in pseudotetrahedral (S_3O) and trigonal planar (S_3) geometries, respectively.[30] These Cd^{II} complexes, $[Cd(\mathbf{TRIL12AL16C})_3]^-$ and $[Cd(\mathbf{TRIL16Pen})_3]^-$, possess very different ^{113}Cd NMR and ^{111m}Cd PAC spectral signatures ($[Cd(\mathbf{TRIL12AL16C})_3]^-$: 574 ppm and $\omega_0 = 0.3405$ rad ns^{-1}; $[Cd(\mathbf{TRIL16Pen})_3]^-$: 684 ppm and $\omega_0 = 0.454$ rad ns^{-1}; Tables 2 and 3). We have proposed that these spectroscopic properties are the result of the different Cd^{II}-coordination numbers and geometries. As support for this model, the parent **TRIL16C** peptide, when interrogated with ^{111m}Cd PAC spectroscopy, appears to bind Cd^{II} as a mixture of both species (with two different signature signals ($\omega_0 = 0.3405$ rad ns^{-1} and $\omega_0 = 0.454$ rad ns^{-1}), while a single coalesced ^{113}Cd NMR resonance (625 ppm) is observed.[25]

Two important observations can be inferred from these data. First, that both species, a pseudotetrahedral CdS_3O and trigonal planar CdS_3, interconvert rapidly on the NMR timescale (0.01–10 ms) but slowly on the PAC timescale (0.1–100 ns).[25,28] The best model to describe this system is a fast water exchange that interconverts the three- and four-coordinate sites quickly. Considering the timescale of both techniques, this water exchange process occurs in the range of ms to ns, a value consistent with the water exchange rates reported for solvated Cd^{II}.[43] Second, this model predicts that there should be a linear correlation between the chemical shift in the ^{113}Cd NMR spectrum and the ratio of CdS_3O/CdS_3 extracted from the ^{111m}Cd PAC spectrum. If the existence of this correlation is true, then the ^{113}Cd NMR chemical shift (which experimentally is more generally available than PAC) could be used as an extremely sensitive indicator of the percentage of four-coordinate CdS_3O and three-coordinate CdS_3 species present in solution.

To assess this last point, we studied a wide range of peptides in the **TRIL16C** family (**TRIL9AL16C**, **TRIL16CL23A** and **TRIL12VL16Pen**) using ^{111m}Cd PAC spectroscopy. These peptides were chosen because their ^{113}Cd NMR chemical shifts (Table 3) predicted either pure CdS_3O or CdS_3 species, or a mixture of both species co-existing. In all cases, the NQIs obtained (Table 2) have values of ω_0 that fall into the two frequency regions obtained

for the pure species $[Cd(\mathbf{TRIL12AL16C})_3]^-$ (ω_0 = 0.3405 rad ns^{-1}) and $Cd(\mathbf{TRIL16Pen})_3]^-$ (ω_0 = 0.454 rad ns^{-1}), indicating that these peptides can in fact be described using the same type of geometries. While the complex $[Cd(\mathbf{TRIL9AL16C})_3]^-$ can be described as a 100% CdS_3O species, the other two peptides, **TRIL16CL23A** and **TRIL12VL16Pen**, bind Cd^{II} as a mixture of CdS_3O and CdS_3. It is worthwhile to note here that the substitution of Leu to Ala above the plane of the Cys seems to selectively stabilize the CdS_3O structure since in both cases, **TRIL12AL16C** and **TRIL9AL16C**, Cd^{II} binds exclusively as a four-coordinate CdS_3O species. Figure 3 shows a plot of the percentage of the four-coordinate species CdS_3O, calculated based on the ^{111m}Cd PAC spectroscopic data (Table 3), against the corresponding ^{113}Cd NMR chemical shift for the different **TRI** peptides. The graphic obtained indicates that indeed there is a linear correlation between both spectroscopic signatures that follows the percentage speciation between the four- and three-coordinate species. This linear correlation predicts a theoretical ^{113}Cd NMR chemical shift of 579 ppm for the four-coordinate CdS_3O species and of 702 ppm for the three-coordinate CdS_3 complex. This last value is very close to the ^{113}Cd NMR chemical shift predicted previously (698 ppm).[30] Our more recent result supports this prediction since the peptide $\mathbf{TRIL12L_dL16C}$ binds Cd^{II} with a 100% trigonal planar geometry, as determined by ^{111m}Cd PAC spectroscopy, and has a ^{113}Cd NMR chemical shift of 697 ppm.[44] Using ^{113}Cd NMR spectroscopy and this linear correlation one can calculate the Cd^{II} speciation present in solution for each of our peptidic systems with very high precision. If one considers that usually ^{111m}Cd PAC spectroscopy can determine the per cent of the different species in solution to within 10% accuracy, a ^{113}Cd NMR chemical shift in the range of 570–600 ppm, and of 675–700 ppm, can be expected for Cd^{II} bound as a CdS_3O and CdS_3, respectively. This ^{113}Cd NMR spectral window is consistent with that observed for similar thiol-rich environments.[14,15]

We know that the **TRIL12AL16C** and **TRIL16Pen** peptides possess a very different pH profile for Cd^{II} binding.[31] As Figure 2 shows, the complex $[Cd(\mathbf{TRIL12AL16C})_3]^-$ is fully formed at pH 7.5 with a pK_{a2} value of 12.2, while a higher pH is required to obtain the complex $[Cd(\mathbf{TRIL16Pen})_3]^-$, which has a pK_{a2} value of 15.8. Therefore, Cd^{II} has a lower pK_{a2} of binding to the four-coordinate site (**TRIL12AL16C**) compared to the three-coordinate site (**TRIL16Pen**). This behavior was also observed in the heterochromic peptide **GRANDL16PenL26AL30C**.[31] Shown in the same Figure 2 there is the pH profile for Cd^{II} binding to the parent peptide **TRIL16C** that binds this metal ion as a mixture of four- and three-coordinate species. This pH profile is intermediate between those obtained for the formation of the complexes $[Cd(\mathbf{TRIL12AL16C})_3]^-$ and $[Cd(\mathbf{TRIL16Pen})_3]^-$, suggesting that the pK_{a2} value for the formation of the different Cd^{II} complexes could also be reflecting the speciation present in solution. If this is the case, and one considers that the ^{113}Cd NMR chemical shift is an indicator of the Cd^{II} speciation as explained above and shown in Figure 3, a linear trend can be expected if one plots the pK_{a2} values against the corresponding ^{113}Cd NMR chemical shifts for the different Cd^{II} complexes of the **TRIL16C** peptide family. Figure 4 demonstrates that this expectation is satisfied. The pK_{a2} value is directly correlated with the ^{113}Cd NMR chemical shifts and, thus, we have another indicator of the percentage of four- and three-coordinate Cd^{II} species, CdS_3O and CdS_3. More important, this correlation provides an independent confirmation of our proposed exchange mechanism.

We next assessed the generality of this correlation by testing whether other designed peptides conformed to the same trends. Two groups of peptides were examined: 1) peptides of the **TRI** peptide family that contain the thiol-rich binding site in a different position from 16, and 2) **GRAND** and **BABY** peptides that are a heptad longer and shorter than **TRI** peptides, respectively. The pK_{a2} values and the corresponding ^{113}Cd NMR chemical shifts of the different **GRAND, BABY** and **TRI** peptides containing the binding site at diverse

positions were determined and included in the plot(Figure 4). The final graphic shows that these systems satisfy the same linear trend and, therefore, we conclude that the ^{113}Cd NMR–pK_{a2} correlation seems to stand independently of the location of the Cys site or the length of the peptide. Furthermore, we have found that the Cd^{II} complex of the thiol derivative of the Coil–Ser peptide, [Cd-(**CSL9C**)$_3$]$^-$, which has a significantly different sequence also follows this linear dependence.[29]

The strategy used in our design to obtain a fully three-coordinate CdS_3 species was the replacement of Leu by the non-natural amino acid penicillamine (Pen), which increases the local bulk directly around the Cd^{II} center. This strategy was successful and leads to peptides with large ^{113}Cd NMR chemical shifts and high pK_{a2} values. While the chemical shift values are a direct representation of the coordination number of the metal, one might expect that the pK_{a2} values could reflect both the coordination number of the metal as well as the chemical properties of the coordinated ligands. Both Cys and Pen have thiol ligands; however, Pen has two methyl groups at the C_β in comparison to two protons for Cys. This substitution is predicted to change the acidity of the thiol group and, thus, potentially the pH of Cd^{II} sequestration. The basicities of both amino acids are similar, pK_a of 8.0 for Pen vs 8.2 for Cys.[45,46] Therefore, the higher pH necessary to bind Cd^{II} to the Pen site would appear to be the result of the intrinsic coordination geometry of the binding site. To verify that this is the case, it was necessary to obtain a fully four-coordinate Cd^{II} binding site using Pen. By analogy with **TRIL12AL16C**, we thought that opening a hole above the Pen layer might generate the desired four-coordinate Cd^{II} and answer this question. A ^{113}Cd NMR chemical shift of 583 ppm was observed at pH 8.5 for the complex [Cd(**TRIL12AL16Pen**)$_3$]$^-$. Based on the ^{113}Cd NMR—^{111m}Cd PAC correlation (Figure 3), this result indicates that we have obtained a four-coordinate CdS_3O species using Pen as a ligand. Conclusive proof for this assignment comes from the ^{111m}Cd PAC data since a single NQI with a ω_0 value of 0.339 rad ns^{-1} was obtained. The results obtained for [Cd(**TRIL12AL16Pen**)$_3$]$^-$ fit nicely in the ^{113}Cd NMR—^{111m}Cd PAC correlation (see Figure 3). This Cd^{II} complex has a higher value of η (0.36) in comparison with the other two complexes containing a 100% CdS_3O, [Cd(**TRIL12AL16C**)$_3$]$^-$ (η = 0.141) and [Cd(**TRIL9AL16C**)$_3$]$^-$ (η = 0.19). The parameter η is the asymmetry parameter and gives information regarding the symmetry of the Cd^{II} center. Its value ranges from 0 to 1, where 0 represents pure axial symmetry. A higher value of η is indicative of a higher degree of distortion in the binding site and, thus, a larger deviation from a tetrahedral geometry. This higher distortion of the four-coordinate Pen site is most likely due to the presence of the methyl groups at the C_β that introduces local hindrance around the thiol donors.

We next determined the pH dependence of Cd^{II} binding to **TRIL12AL16Pen**. As shown in Figure 2 (open circles) the pH profile obtained reveals that the coordination geometry, and not the ligand basicity, is the factor controlling the site acidity. The pK_{a2} value drops from 15.8 to 12.7 in moving from a three-coordinate to a four-coordinate Pen metal-binding site. Consistent with these observations, a pK_{a2} of 15.1 is obtained for Cd^{II} binding to **TRIL12L$_D$L16C**, a peptide that binds Cd^{II} as a fully 100% CdS_3 and contains Cys as a coordinating ligand.[44]

Figure 4 demonstrates that all four-coordinate sites, independently of the nature of the thiol group (Cys or Pen), bind Cd^{II} fully under lower pH conditions than do the three-coordinate sites. Another feature is apparent upon closer examination of this correlation. The pK_{a2} values obtained for peptides that bind Cd^{II} exclusively as three-coordinate CdS_3 species have little variation, whereas those peptides sequestering Cd^{II} as a four-coordinate CdS_3O species exhibit a wide range for the pK_{a2} values (11.3 to 13.3). These data indicate that moving from **TRI** to **GRAND** peptides or by introducing double substitutions, the pH dependence of Cd^{II} binding to the three-coordinate site is essentially unaffected, while these

modifications significantly perturb the four-coordinate site. This variation most likely is a reflection of different self-association affinities and pH-dependent aggregation state preferences of the peptides themselves. For example, the **GRAND** peptides show a higher self-association affinity than **TRI** peptides, ≈5 kcal mol^{-1} per coiled coil aggregate. On the other hand, the disruption of a Leu layer by smaller nonpolar amino acids, such as Val and Ala, can destabilize the final coiled coil up to 6–7 kcal mol^{-1}, reflecting poorer hydrophobic packing.[26] The greater stability of the **GRAND** peptides result in the formation of three-stranded coiled coils at pH conditions lower than 6.0. This enhanced aggregate stability provides a scaffold which can complex Cd^{II} at lower pH values than the corresponding **TRI** peptides. At the other extreme, coiled coils that have been destabilized by disrupting the core packing require higher pH to form the assembly which is necessary to sequester the metal properly. In support of this model, we observe that an increase in the concentration of the **TRI** peptide with respect to the amount of Cd^{II} shifts the pH profile for formation of the complex to lower pH values. This suggests that the formation/stabilization of the four-coordinate Cd^{II} may be achieved at more acidic pH values than those observed here by increasing the self-association energy of the three-stranded coiled coil. This explanation also rationalizes the small variance in pK_{a2} values for three-coordinate CdS_3 peptides containing penicillamine. The Pen substitution stabilizes the three-stranded coiled coil significantly and, since high pH is required to form the trigonal structure, these peptides all exhibit high self-affinity under these conditions. The data indicate that the lower limit for the pK_{a2} value of this type of four-coordinate site is determined in our systems by the actual formation of the binding site. The data included in the supporting information corroborates all these observations.

This behavior is very different from that observed for the Hg^{II} complexes of the **TRI** peptide family where the formation of a three-coordinate thiolate Hg^{II} center does not depend on the stability of the coiled coil.[47–49] In this case, there are two relevant equilibria. The first is the conversion of a two-stranded coiled coil to a three-stranded coiled coil aggregate. This equilibrium appears to be metal independent having a pK_a for **TRI**, with and without Hg^{II}, of approximately 5.5. The second equilibrium is between a three-stranded coiled coil containing linear Hg^{II} as HgS_2 and the trigonal planar Hg^{II} species HgS_3. As long as one examines peptides that have Cys substitution in the "**a**" heptad position, one observes the conversion of Hg-$(peptide)_2$(peptide−H) to $Hg(peptide)_3^-$ plus loss of a proton with a pK_a = 7.7. This is true regardless of whether the peptide used is **TRIL16C**, **TRIL9C**, **BABYL9C** or **GRANDL9C**.[47, 48]

At this point, we have obtained two important correlations between CdS_3O centers ($\omega_0 \approx$ 0.34 rad ns^{-1}) and CdS_3 centers ($\omega_0 \approx 0.45$ rad ns^{-1}) that allow the determination, based only in the ^{113}Cd NMR chemical shift, of the percentage of four-coordinate CdS_3O and three-coordinate CdS_3 species (Figure 3) and the pH required for full Cd^{II} binding (Figure 4). Both correlations track the ratio of both species in solution. Often one sees quoted in references that a specific coordination geometry (e.g., CdS_3O) has a broad ^{113}Cd NMR chemical shift range without an explanation for the factors that cause the large variation. To our knowledge, this study provides the first systematic explanation for why the ^{113}Cd NMR chemical shift range can vary over what might be perceived to be 125 ppm. Furthermore, we have demonstrated for the first time that ^{113}Cd NMR can be used as a highly sensitive probe of the equilibrium contributions of two rapidly interchanging protein bound species when the two limiting species have been clearly identified. The next question that we must ask is how general this correlation is.

To address this point we must recognize that a four-coordinate Cd^{II} with three endogenous sulfur ligands can vary its structure in several ways within our designed peptides. First, is the issue of metal binding to "**a**" versus "**d**" type sites. Both Cd^{II} and Hg^{II} show, for

example, different pK_a values for formation of MS_3 species when the site is in an "**a**" versus "**d**" position. Modelling studies suggest that the orientation of Cys in these two sites is markedly different with Cys sulfurs in an "**a**" site positioned towards the central axis of the helix whereas "**d**" sites orient the Cys towards the helical interfaces. Second, the position of the metal binding site may be displaced either towards the N- or C-terminal ends of the peptides. As an example of this discrepancy we can consider our recent crystallographic study of As-(**CSL9C**)$_3$, an analogous peptidic system to **TRI**, that has provided the first X-ray structure of As^{III} coordinated to three Cys residues in a protein environment (Figure 5A).[50] The As^{III} was found to be in a trigonal pyramidal geometry with As–S bond lengths (2.28 ± 0.05 Å) and S-As-S angles (92°) in good agreement with small molecule complexes and the EXAFS distances previously determined for As^{III} coordination to **TRIL12C**, **TRIL16C**, ArsR and ArsD.[51–54] Most important, the As^{III} ion was found to be coordinated below the plane of the β-methylene protons of Cys in an *endo* conformation and not in an *exo* conformation capping the Cys residues as was previously assumed by most workers in this field.[55] On the other hand, Pb^{II}, which also prefers a trigonal pyramidal thiol-rich binding site, shows an *exo* configuration being coordinated above the thiolate plane when bound to the protein aminolevulinic acid dehydratase (Figure 5B).[56] This binding mode is consistent with the larger size of Pb^{II} in comparison to As^{III}, with typical Pb^{II}–S distances of 2.62–2.67 Å for three-coordinate PbS_3 structures.[33, 57–60] Consistent with these data, our EXAFS spectroscopic studies demonstrate that Pb^{II} binds to **TRIL16C** and **TRIL12C** peptides as a PbS_3 structure with Pb^{II}–S distances of 2.63 Å.[27] Cd^{II} is intermediate in size between As^{III} and Pb^{II} and has the potential to bind to peptides in either the *endo* or *exo* conformations as is demonstrated in Figure 5C and D. However, unlike both As^{III} and Pb^{II} which contain stereochemically active lone pairs filling the fourth coordination sites, Cd^{II} can bind a water molecule in this fourth position.

From the ^{113}Cd NMR and ^{111m}Cd PAC spectroscopic data we know that the **TRIL16C** peptide binds Cd^{II} as a mixture of four- and three-coordinate Cd^{II} species (CdS_3O and CdS_3).[25] Certainly, the CdS_3 polyhedron is trigonal planar; however, one must consider whether the Cd^{II} in the CdS_3O structure is in an *exo* or *endo* conformation. The replacement of the Leu layer above the metal thiol-rich binding site by Ala has proven sufficient to generate two peptides, **TRIL12AL16C**[30] and **TRIL12AL16Pen**, capable of binding Cd^{II} solely with a four-coordinate CdS_3O distorted tetrahedral geometry. This mutation should generate sufficient space above the thiol ligands to allow water access to the hydrophobic interior. This observation would suggest that Cd^{II} should bind to these sites in a way that resembles most closely the Pb^{II} *exo* conformation rather than the As^{III} *endo*. Such an assignment is consistent with an NMR structure of Cd^{II} bound to the HIV-I integrase H12C mutant.[61] The observation that when the second Leu layer above the Cys is replaced with Ala (**TRIL9AL16C**) the peptide binds Cd^{II} as solely a CdS_3O species, whereas the equivalent mutation below the Cys layer (**TRIL16CL23A**) has essentially no effect on the ratio of CdS_3O and CdS_3 as compared to **TRIL16C**, further supports an *exo* conformation. Our best interpretation of these results is that opening a hole at position 9 could affect the hydrophobic packing of the Leu at position 12 allowing the entrance of water to the hydrophobic interior and thus its coordination to Cd^{II}. It should be noted that these three peptides (**TRIL12AL16C**, **TRIL12AL16Pen** and **TRIL9AL16C**) all exhibit PAC spectra with $\omega_0 \approx 0.34$ rad ns^{-1}. In fact, every peptide complex reported in the correlations of Figure 3 and 4 which contain a CdS_3O species also shows ω_0 ~0.34 rad ns^{-1}. Thus, we conclude that the correlations that we have discovered are directly applicable to understanding equilibria between CdS_3 and CdS_3O species bound in "**a**" site peptides with an *exo* metal conformation.

One might then ask what are the properties of CdS_3O sites oriented in an *endo* conformation or bound within a "**d**" site environment. Unfortunately, sufficient data do not yet exist for

such constructs to evaluate these possibilities fully; however, we have been able to isolate one system that may begin to address one of these questions. The peptide **TRIL16CL19A** engineers a water-binding pocket below the metal binding site which might induce a shift in metal environment to the *endo* conformation. The ^{113}Cd NMR chemical shift for [Cd(**TRIL16CL19A**)$_3$]$^-$ is 605 ppm and the PAC spectrum indicates that this peptide contains 100% CdS_3O but now with an $\omega_0 = 0.274$ rad ns^{-1}. This chemical shift value and composition is inconsistent with the trend found in Figure 3. Furthermore, the pK_{a2} for this peptide is 11.2 a value which again is inconsistent with the 16 points reported in Figure 4. The tempting conclusion is that CdS_3O sites with $\omega_0 > 0.33$ rad ns^{-1} correspond to *exo* conformations of the Cd^{II} whereas $\omega_0 < 0.28$ rad ns^{-1} may indicate the *endo* conformers. Furthermore, a change in the *exo* versus *endo* conformation may lead to different trend lines for the scaling of the CdS_3O/CdS_3 ratio with NMR and the acidity of the corresponding metal binding site. These intriguing possibilities must await further study to be confirmed; however, the concepts presented by these data raise interesting general questions for Cd^{II} binding to proteins, the use of Cd^{II} as a probe for Zn^{II} and the interpretation of ^{113}Cd NMR for Cd^{II}-substituted Zn^{II} sites in biology.

The pioneers investigating Cd^{II} binding to proteins, especially as a spectroscopically amenable substitute for Zn^{II}, were confined to those systems that became available and which may or may not have been structurally characterized. This meant that broad chemical shift ranges could be mapped; however, detailed questions relating to specific geometries or the impact of protein dynamics were difficult to explore. In this article, we have demonstrated how the conjunction of protein design and ^{111m}Cd PAC spectroscopy allows for a more refined level of understanding for biological ^{113}Cd NMR. The systematic variation of protein constructs using de novo designed peptides allows the control of most structural variables for a metalloprotein. Thus, one may make a subtle sequence change that perturbs a highly defined metal chromophore in an interpretable manner. In this way it becomes straightforward to prepare a suite of proteins that can be used to interrogate defined physical properties of interest (e.g., ^{113}Cd NMR shift or pK_a). PAC spectroscopy provides a faster timescale to interrogate the metal center. In this case, it revealed that what may have been interpreted as a single species based on ^{113}Cd NMR was in fact an equilibrium between two fast exchanging species. Furthermore, by extracting the NQI parameters, one may obtain useful information about the metal site such as whether a pseudotetrahedral metal binds in an *endo* or an *exo* conformation. Therefore, one is now able to systematically vary the properties of well defined binding sites so that correlations as presented here may be deduced.

The first of these correlations provides decisive proof for the CdS_3 to CdS_3O equilibrium model and demonstrates that one may use ^{113}Cd NMR as an accurate and convenient method to assess the ratios of these two species. Moreover, it illustrates that Cd^{II} bound to a protein with fast ligand exchange can exhibit intermediary physical properties between the two exchanging species. Certainly this is true for the coalesced ^{113}Cd NMR chemical shift, but more importantly, the pH at which the complex converts to having all sulfurs bound to the metal (either as CdS_3O or CdS_3) is dependent upon the ratio of these two components. For this system, if one knows the ^{113}Cd chemical shift, then one can predict with a high degree of accuracy the pK_{a2} for complex formation. Since the binding affinity of the metal to the protein is enhanced when all three sulfurs bind, this suggests that the thermodynamically more stable constructs will be enriched in the CdS_3O component. Cd^{II} binding to heterochromic polypeptides exhibits just this specificity and selectivity.[31]

To date, there has been the tendency to assign a ^{113}Cd chemical shift to a single, specific coordination structure and to support this assignment by indicating that the chemical shift value falls within a broad range of shifts consistent with the assigned structure. Our studies

indicate that one should consider a dynamic model for metal binding which may include more than one species. Furthermore, these observations suggest that a protein may be able to fine tune the affinity of Cd^{II} binding by developing a metal site that controls the ratio of CdS_3O to CdS_3 (or in principle two other species of different affinity which are in fast exchange).

One example where such a process may be operable is the SmtB type metalloregulatory protein CadC.[20] Site-directed mutagenesis and Cd EXAFS spectroscopy support the assignment of the Cd^{II} binding site as CdS_4. The ^{113}Cd chemical shift is 625 ppm, a value that is comfortably within literature values for CdS_4 (typically reported as 590 to 750 ppm).[12, 15] However, it is surprising that the CdS_4 range is so large and overlaps so significantly with CdS_3-type complexes. Typically, the addition of each sulphur to the Cd^{II}-coordination sphere causes a significant downfield shift, yet CdS_4 overlaps CdS_3 shifts for over half their range (from 590 to 700 ppm). We have noticed that the pure species CdS_3 or CdS_3O have relatively narrow chemical shift ranges. With pure CdS_3 species, we observe a range of chemical shifts of ≈25 ppm (from 675 to 700 ppm). The same can be said for CdS_3O which appears to span 570–600 ppm. The 125 ppm breadth that occurs is a result of the equilibrium between these two species. It is tempting to suggest that the even broader range reported for CdS_4 (160 ppm) could also be a consequence of mixed sites. If this reasoning is correct, then a more appropriate value for a pure CdS_4 species would likely be closer to 750 ppm, with a lower limit of ≈720 ppm. If a pure CdS_4 in a protein environment truly has this chemical shift range, how then could one explain the 625 ppm resonance of CadC? The simplest explanation would be that Cd^{II} might bind as a mixture of CdS_4 and CdS_3O, with water as the fourth ligand. This makes the CadC system an ideal target for PAC studies to test this hypothesis.

The situation becomes even more interesting when considering proteins designed to bind Zn^{II}. Important examples for our studies are the ZnS_3O site of hepatitis C virus NS3 proteinase and an HIV-1 integrase mutant which also forms a ZnS_3O structure.[22, 61] While it has become commonplace to substitute Cd^{II} for Zn^{II} in proteins like these, the replacement is not completely benign. Cd^{II} is larger and has longer bond lengths than Zn^{II} (e.g., for CdS_3O the Cd–S bond is approximately 2.54 Å whereas the corresponding Zn^{II} distance is 2.32 Å).[42, 62] Furthermore, Cd^{II} is somewhat softer than Zn^{II} and prone to form higher coordination number complexes. In the case of substitution into an identical protein environment the most likely concern is size. Size can shift the preference of a metal from an *exo* to an *endo* conformation as has been reported for As(**CSL9C**)$_3$.[50] The differential size may also restrict the ability for an exogenous ligand such as water to bind to the metal. Thus, in metal-binding sites such as those found in the **TRI** peptides, steric constraints appear to favour the formation of a significant percentage of CdS_3 centers (e.g., **TRIL16C**). The 0.2 Å shorter bond lengths of ZnS_3O should not be as influenced by this steric constraint, suggesting that Zn^{II} might bind to all, or most, of these designed peptides as a four coordinate complex with bound water. What these data would suggest is that in cases where an equilibrium process is in effect for Cd^{II} in a protein, cadmium substitution may not provide fully appropriate information about the zinc environment.

We have proposed a number of tantalizing questions which may have profound impact for understanding biological cadmium coordination chemistry, the interpretation of ^{113}Cd NMR spectra and the application of Cd^{II} substitution as a reporter for the highly prevalent Zn^{II} class of metalloproteins. However, there are some caveats that must be considered at this stage. While we have a reasonable explanation for the large chemical shift range for at least one coordination environment (ostensibly CdS_3O) and have enumerated several factors that contribute to the chemical shift variation, our correlation is, in fact, limited to CdS_3 trigonal planar to CdS_3O pseudotetrahedral structures found in "**a**" type sites, presumably with the

metal in an *exo* configuration. One might expect some variation to our reported trends if the metal is located in "**d**" sites or in an *endo* conformation. Furthermore, other factors such as the change in orientation of thiolate donors on converting from parallel to antiparallel helices, and hydrogen bonding could further broaden the chemical shift ranges discussed here. Because of the importance of zinc metalloproteins and the applications of Cd^{II} for understanding the metal center structure and function, it is clear that additional studies addressing each of these points is essential. We have shown how de novo designed peptides can be powerful vehicles to examine these important issues. We expect that future studies with variants of these peptides will further elucidate this fascinating and important problem.

Experimental Section

Peptide synthesis and purification

The **TRI** peptides and derivatives (see Table 1 for sequence nomenclature) were synthesized on an Applied Biosystems 433 A peptide synthesizer using standard F-moc protocols,[63] and purified and characterized as described previously.[47]

PAC spectroscopy

All perturbed angular correlation (PAC) experiments were performed with a setup using six detectors and a temperature of 1 ± 2 °C, which was controlled by a Peltier element. The radioactive cadmium was produced on the day of the experiment at the University Hospital cyclotron in Copenhagen and extracted as described previously,[64] except for the HPLC separation of zinc and cadmium which was omitted. This procedure may lead to zinc contamination of the sample, but the level of contamination should not interfere with the experiment. The ^{111m}Cd solution (10–40 μL) was mixed with nonradioactive cadmium acetate and TRIS buffer. The peptide, dissolved in ion-exchanged water, was then added 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. In order to avoid chloride contamination, a small volume of sample was removed from the solution and the pH measured. The pH reported in Table 2 was measured at room temperature the following day and corrected to the pH at 1 °C as follows. The pH of solutions buffered by TRIS is temperature dependent; therefore the pH of the solutions at 1 °C was calculated using the following equation pH(1 °C) 0.964[pH(25 °C)] + 0.86.[65] 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 μ_M peptide, 20 m_M TRIS buffer and a Cd^{II}/peptide ratio of 1:12. All fits were carried out with 400 data points, disregarding the 3–5 first points due to systematic errors in these. For the **TRIL12AL16Pen** peptide, 300 points were used in the fit.

Each nuclear quadrupole interaction (NQI) was modelled using a separate set of parameters that includes ω_0, η, $\Delta\omega_0/\omega_0$, $1/\tau_c$ and A. The parameter ω_0 [$\omega_0 = 12\pi \mid eQV_{zz} \mid /(40\ h)$, where Q is the nuclear electric quadrupole moment and V_{zz} is the numerically largest eigenvalue of the electric field gradient tensor] is associated with the strength of the interaction between the surrounding electronic environment and the Cd nucleus, η is the so-called asymmetry parameter which is 0 in an axially symmetric complex and has a maximal value of 1; $\Delta\omega_0/\omega_0$ describes static structural variations from one Cd^{II} site to the next, and is as such a measure of the structural variability; τ_c is the rotational correlation time; and A is the amplitude of the signal (see ref. [25] for a more detailed description). The parameters fitted to the PAC data are presented in Table 2. The percentage of each species reported in Table 3 were calculated based on the A value.

^{113}Cd NMR spectroscopy

All the spectra were collected at room temperature on a Varian Inova 500 spectrometer (110.92 MHz for ^{113}Cd) equipped with a 5 mm broadband probe. ^{113}Cd NMR spectra were externally referenced to a 0.1 M $Cd(ClO_4)_2$ solution in D_2O. A spectral width of 847 ppm (93 897 Hz) was sampled using a 5.0 μs 90° pulse and 0.05 s acquisition time with no delay between scans. Samples were prepared under a flow of argon or nitrogen by dissolving 30–35 mg of the lyophilized and degassed peptides in 450–500 μL 15% D_2O solution. The peptide concentrations were determined by using either the Ellman's[66] or 4,4′-dipyridyl disulphide tests,[67] and the concentrations range from 12 to 18 mM peptide, which corresponds to 4–6 mM three-stranded coiled coil. The final samples were prepared by the addition of the appropriate amount of 250 mM $^{113}Cd(NO_3)_2$ solution (prepared from 95% isotopically enriched ^{113}CdO obtained from Oak Ridge National Laboratory) and the adjustment of the pH with KOH or HCl solutions. The final pH value for each peptide was chosen based on its pH titration curves. This pH value correspond to the fully formation of the Cd^{II} complex. An Argon or Nitrogen atmosphere was maintained when possible but the samples came in contact with O_2 while the pH was adjusted. The actual final concentrations for each experiment are indicated in the text and the Figure captions. The data were analyzed using the software MestRe-C.[68] All free induction decays (FID's) were zero filled to double the original points and were processed by application of 50 Hz line broadening prior to Fourier transformation.

UV/Vis spectroscopy

All the UV/Vis pH titration experiments were carried out at room temperature on an Ocean Optics SD 2000 fiber optic spectrometer. Fresh stock solutions of the purified peptides were prepared for each experiment using doubly distilled water and were purged with argon to minimize the chances of oxidation. The peptide concentration was determined as described in the ^{113}Cd NMR section.[66, 67] pH Titrations were performed by adding small aliquots of KOH (1 mM to 1 M stock solutions) to unbuffered solutions containing $CdCl_2$ salt (20 μM) and peptide (60–120 μM). All the doubly distilled water used in these experiments was bubbled with Ar before addition of $CdCl_2$ and peptide. The change in pH was monitored using an Accumet gel-filled pencil-thin Ag/AgCl single-junction electrode with an Orion Research digital pH millivolt meter 611. In all cases, equilibration time was allowed before reading the final pH and it was found that binding of Cd^{II} to **TRIL16Pen**, **TRIL12VL16Pen** and **GrandL16Pen** was slower than binding the other peptides. To verify the reversibility of the process, reverse titrations were carried out subsequently in all the experiments by adding small aliquots of 1 mM to 1 M solutions of HCl. The pH-dependent absorption spectra were fit using the same models and procedures as those used in our previous studies.[27, 29] Specifically, the experimental data were analyzed by non-linear least square fitting to the following equation:

$$Cd[(peptide{-}H)_2(peptide)]^+ = Cd(peptide)_3^- + 2H^+ \quad Abs = (c[M]_t * \Delta\varepsilon)/(10^{(pK_{a2}-2pH)}+1)$$

where Abs is the observed absorbance after addition of KOH, $c[M]_t$ is the total metal concentration, $\Delta\varepsilon$ is the extinction coefficient of the metal complex, and pK_{a2} the acidity constant for the simultaneous release of two protons. The data were also fit using an extended model including additional species in solution (see Supporting Information for a full description).

Acknowledgments

V.L.P. thanks the National Institute of Health for support of this research (R01 ES0 12236), O.I. thanks the Margaret and Herman Sokol Foundation for a Postdoctoral Award and L.H. thanks The Danish Natural Science Research Council for support.

References

1. Vallee BL, Auld DS. Biochemistry. 1990; 29:5647–5659. [PubMed: 2200508]
2. Lipscomb WN, Sträter N. Chem. Rev. 1996; 96:2375–2434. [PubMed: 11848831]
3. Coleman JE. Curr. Opin. Chem. Biol. 1998; 2:222–234. [PubMed: 9667939]
4. Berg JM, Shi Y. Science. 1996; 271:1081–1085. [PubMed: 8599083]
5. Klug A. J. Mol. Biol. 1999; 293:215–218. [PubMed: 10529348]
6. McCall KA, Huang CC, Fierke CA. J. Nutr. 2000; 130:1437S–1446S. [PubMed: 10801957]
7. Tobin DA, Pickett JS, Hartman HL, Fierke CA, Penner-Hahn JE. J. Am. Chem. Soc. 2003; 125:9962–9969. [PubMed: 12914459]
8. Sousa SF, Fernandes PA, Ramos MJ. J. Biol. Inorg. Chem. 2005; 10:3–10. [PubMed: 15611883]
9. Myers LC, Terranova MP, Nash HM, Markus MA, Verdine GL. Biochemistry. 1992; 31:4541–4547. [PubMed: 1581309]
10. Penner-Hahn J. Curr. Opin. Chem. Biol. 2007; 11:166–171. [PubMed: 17376731]
11. Andreini C, Banchi L, Bertini I, Rosato A. J. Proteome Res. 2006; 5:196–201. [PubMed: 16396512]
12. Ecoleman J. Methods Enzymol. 1993; 227:16–43. [PubMed: 8255225]
13. Maret W, Vallee BL. Methods Enzymol. 1993; 226:52–71. [PubMed: 8277880]
14. Summers MF. Coord. Chem. Rev. 1988; 86:43–134.
15. Öz G, Pountney DL, Armitage IM. Biochem. Cell Biol. 1998; 76:223–243. [PubMed: 9923691]
16. Hemmingsen L, Nárcisz K, Danielsen E. Chem. Rev. 2004; 104:4027–4061. [PubMed: 15352785]
17. Hemmingsen L, Damblon C, Antony J, Jensen M, Adolph HW, Wommer S, Roberts GCK, Bauer R. J. Am. Chem. Soc. 2001; 123:10329–10335. [PubMed: 11603983]
18. Kofod P, Bauer R, Danielsen E, Larsen E, Bjerrum MJ. Eur. J. Biochem. 1991; 198:607–611. [PubMed: 2050141]
19. Bauer R, Bjerrum MJ, Danielsen E, Kofod P. Acta Chem. Scand. 1991; 45:593–603. [PubMed: 1764331]
20. Busenlehner LS, Cosper NJ, Scott RA, Rosen BP, Wong MD, Giedroc DP. Biochemistry. 2001; 40:4426–4436. [PubMed: 11284699]
21. Erskine PT, Norton E, Cooper JB, Lambert R, Coker A, Lewis G, Spencer P, Sarwar M, Wood SP, Warren MJ, Shoolingin-Jordan PM. Biochemistry. 1999; 38:4266–4276. [PubMed: 10194344]
22. Stempniak M, Hostomska Z, Nodes BR, Hostomsky Z. J. Virol. 1997; 71:2881–2886. [PubMed: 9060645]
23. 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.
24. a) Dieckmann GR, McRorie DK, Lear JD, Sharp KA, DeGrado WF, Pecoraro VL. J. Mol. Biol. 1998; 280:897–912. [PubMed: 9671558] b) Peacock AFA, Iranzo O, Pecoraro VL. Dalton Trans. 2009
25. Matzapetakis M, Farrer BT, Weng T-C, Hemmingsen L, Penner-Hahn JE, Pecoraro VL. J. Am. Chem. Soc. 2002; 124:8042–8054. [PubMed: 12095348]
26. Ghosh D, Lee K-H, Demeler B, Pecoraro VL. Biochemistry. 2005; 44:10732–10740. [PubMed: 16060682]
27. Matzapetakis M, Ghosh D, Weng K-H, Penner-Hahn JE, Pecoraro VL. J. Biol. Inorg. Chem. 2006; 11:876–890. [PubMed: 16855818]
28. Lee K-H, Matzapetakis M, Mitra S, Marsh ENG, Pecoraro VL. J. Am. Chem. Soc. 2004; 126:9178–9179. [PubMed: 15281796]
29. Iranzo O, Ghosh D, Pecoraro VL. Inorg. Chem. 2006; 45:9959–9973. [PubMed: 17140192]
30. Lee K-H, Cabello C, Hemmingsen L, Marsh ENG, Pecoraro VL. Angew. Chem. 2006; 118:2930–2934. *Angew. Chem. Int. Ed.***2006**, *45*, 2864 –2868.
31. Iranzo O, Cabello C, Pecoraro VL. Angew. Chem. 2007; 119:6808–6811. *Angew. Chem. Int. Ed.* **2007**, *46*, 6688 –6691.
32. Matthews RG, Goulding CW. Curr. Opin. Chem. Biol. 1997; 1:332–339. [PubMed: 9667865]

33. Erskine PT, Senior N, Awan S, Lambert R, Lewis G, Tickle IJ, Sarwar M, Spencer P, Thomas P, Warren MJ, Shoolingin-Jordan PM, Wood SP, Cooper JB. Nat. Struct. Biol. 1997; 4:1025–1031. [PubMed: 9406553]
34. González B, Pajares MA, Martínez-Ripoll M, Blundell TL, Sanz-Aparicio J. J. Mol. Biol. 2004; 338:771–782. [PubMed: 15099744]
35. Kim JL, Morgenstern KA, Lin1 TFC, Dwyer MD, Landro JA, Chambers SP, Markland W, Lepre CA, O'Malley ET, Harbeson SL, Rice CM, Murcko MA, Caron PR, Thomson JA. Cell. 1996; 87:343–355. [PubMed: 8861917]
36. Rice WG, Supko JG, Malspeis L, Clanton RWBD Jr, Bu M, Graham L, Schaeffer CA, Turpin JA, Domagala J, Gogliotti R, Bader JP, Halliday SM, Coren L, S RC II, Arthur LO, Henderson LE. Science. 1995; 270:1194–1197. [PubMed: 7502043]
37. Cai M, Zheng R, Caffrey M, Craigie R, Clore GM, Gronenborn AM. Nat. Struct. Biol. 1997; 4:567–577. [PubMed: 9228950]
38. Santos RA, Gruff ES, Koch SA, Harbison GS. J. Am. Chem. Soc. 1991; 113:469–475.
39. Alberts IL, Nadassy K, Wodak SJ. Protein Sci. 1998; 7:1700–1716. [PubMed: 10082367]
40. Coyne HJ III, Ciofi-Baffoni S, Banci L, Bertini I, Zhang L, George GM, Winge DR. J. Biol. Chem. 2007; 282:8926–8934. [PubMed: 17215247]
41. Castro C, Millian NS, Garrow TA. Arch. Biochem. Biophys. 2008; 472:26–33. [PubMed: 18262489]
42. Parkin G. Chem. Rev. 2004; 104:699–767. [PubMed: 14871139]
43. Docommun, Y.; Merbach, AE., et al. Inorganic High Pressure Chemistry: Kinetics and Mechanisms. Eldik, Rv, editor. Amsterdam: Elsevier; 1986. p. 70
44. Peacock AFA, Hemmingsen L, Pecoraro VL. Proc. Natl. Acad. Sci. USA. 2008; 105:16566–16571. [PubMed: 18940928]
45. Boggess RK, Absher JR, Morelen S, Taylor LT, Hughes JW. Inorg. Chem. 1983; 22:1273–1279.
46. Cheesman BV, Arnold AP, Rabenstein DL. J. Am. Chem. Soc. 1988; 110:6359–6364.
47. Farrer BT, Harris NP, Balchus KE, Pecoraro VL. Biochemistry. 2001; 40:14696–14705. [PubMed: 11724584]
48. Ghosh D, Pecoraro VL. Inorg. Chem. 2004; 43:7902–7915. [PubMed: 15578824]
49. Iranzo O, Thulstrup PV, Ryu S-B, Hemmingsen L, Pecoraro VL. Chem. Eur. J. 2007; 13:9178–9190. [PubMed: 17960740]
50. Touw DS, Nordman CE, Stuckey JA, Pecoraro VL. Proc. Natl. Acad. Sci. USA. 2007; 104:11969–11974. [PubMed: 17609383]
51. Farrer BT, McClure CP, Penner-Hahn JE, Pecoraro VL. Inorg. Chem. 2000; 39:5422–5423. [PubMed: 11154553]
52. Shaikh TA, Bakus RC II, Parkin S, Atwood DA. J. Organomet. Chem. 2006; 691:1825–1833.
53. Shaikh TA, Parkin S, Atwood DA. J. Organomet. Chem. 2006; 691:4167.
54. Shi W, Dong J, Scott RA, Ksenzenko MY, Rosen BP. J. Biol. Chem. 1996; 271:9291–9297. [PubMed: 8621591]
55. Carter TG, Healey ER, Pitt MA, Johnson DW. Inorg. Chem. 2005; 44:9634–9636. [PubMed: 16363829]
56. Erskine PT, Duke EMH, Tickle IJ, Senior NM, Warren MJ, Cooper JB. Acta Crystallogr. Sect. D. 2000; 56:421–430. [PubMed: 10739915]
57. Dean PAW, Vittal JJ, Payne NC. Inorg. Chem. 1984; 23:4232–4236.
58. Christou G, Folting K, Huffman JC. Polyhedron. 1984; 3:1247–1253.
59. Magyar JS, Weng T-C, Stern CM, Dye DF, Rous BW, Payne JC, Bridgewater BM, Mijovilovich A, Parkin G, Zaleski JM, Penner-Hahn JE, Godwin HA. J. Am. Chem. Soc. 2005; 127:9495–9505. [PubMed: 15984876]
60. Busenlehner LS, Weng T-C, Penner-Hahn JE, Giedroc DP. J. Mol. Biol. 2002; 319:685–701. [PubMed: 12054863]
61. Cai M, Huang Y, Caffrey M, Zheng R, Craigie R, Clore GM, Gronenborn AM. Protein Sci. 1998; 7:2669–2674. [PubMed: 9865962]

62. Govindaswamy N, Moy J, Millar M, Koch SA. Inorg. Chem. 1992; 31:5343–5344.
63. Chan, WC.; White, PD. Fmoc Solid Phase Peptide Synthesis: A Practical Approach. New York: Oxford University Press; 2000.
64. Hemmingsen L, Bauer R, Bjerrum MJ, Zeppezauer M, Adolph HW, Formicka G, Cedergren-Zeppezauer E. Biochemistry. 1995; 34:7145–7153. [PubMed: 7766625]
65. Hemmingsen LB, Bjerrum R, Adolph M, Zeppezauer H, Cedergren-Zeppezauer ME. Eur. J. Biochem. 1996; 241:546–551. [PubMed: 8917454]
66. Ellman GL. Arch. Biochem. Biophys. 1959; 82:70–77. [PubMed: 13650640]
67. Mantle M, Stewart G, Zayas G, King M. Biochem. J. 1990; 266:597–604. [PubMed: 2317206]
68. Cobas, C.; Cruces, J.; Sardina, FJ. MestRe-C Version 2.3. Spain: Universidad de Santiago de Compostela; 2000.

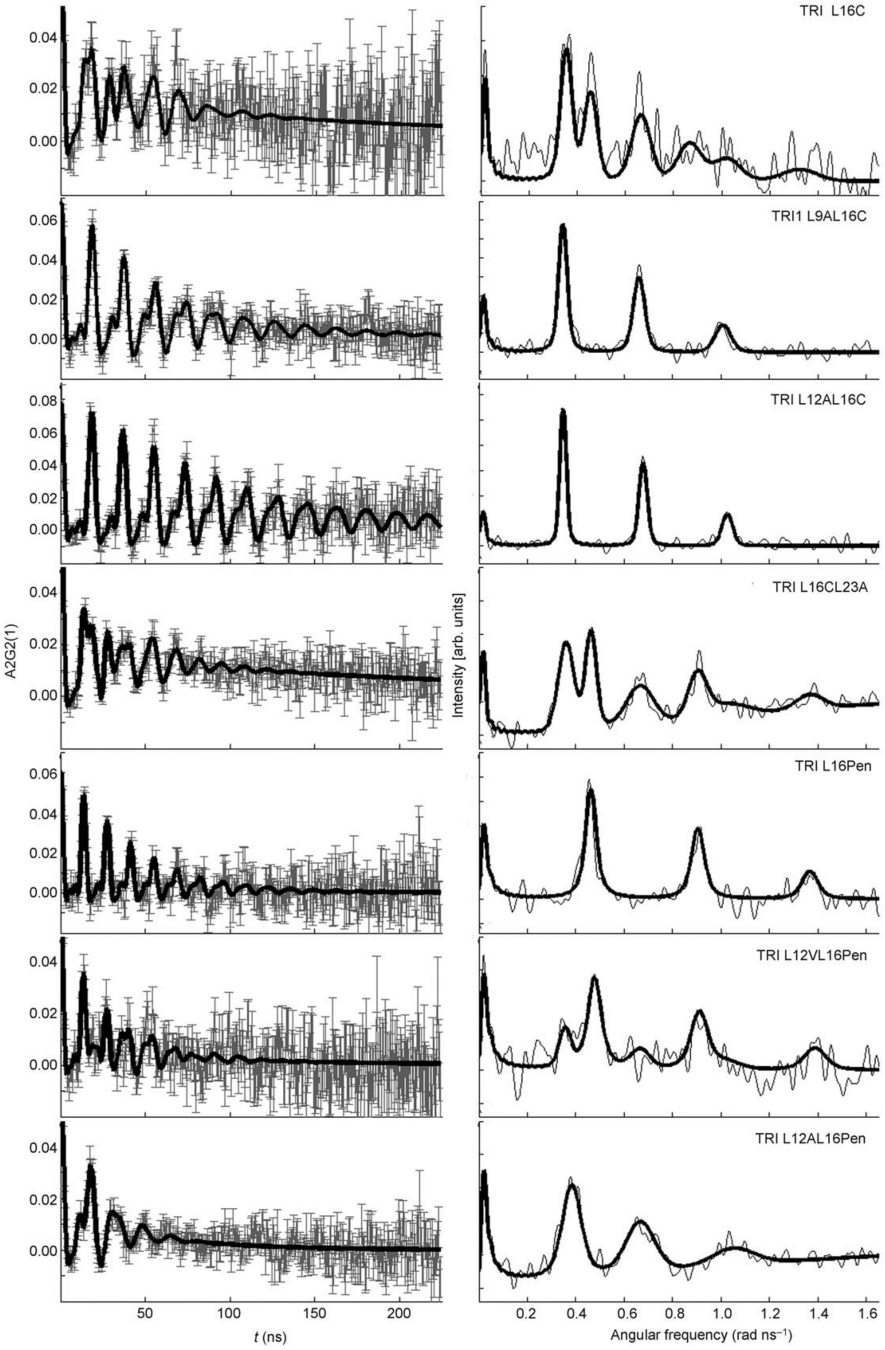

Figure 1.

^{111m}Cd PAC spectra of the different **TRI** peptides. All the samples contained 20 mM TRIS buffer, 300 μM peptide and a Cd^{II}/peptide ratio of 1:12. Left: Perturbation function (experimental data with error bars are shown in grey, fit shown as solid line); right: the Fourier transform (the thin line gives the Fourier transform of the experimental data, and the bold faced line gives the Fourier transform of the fit).

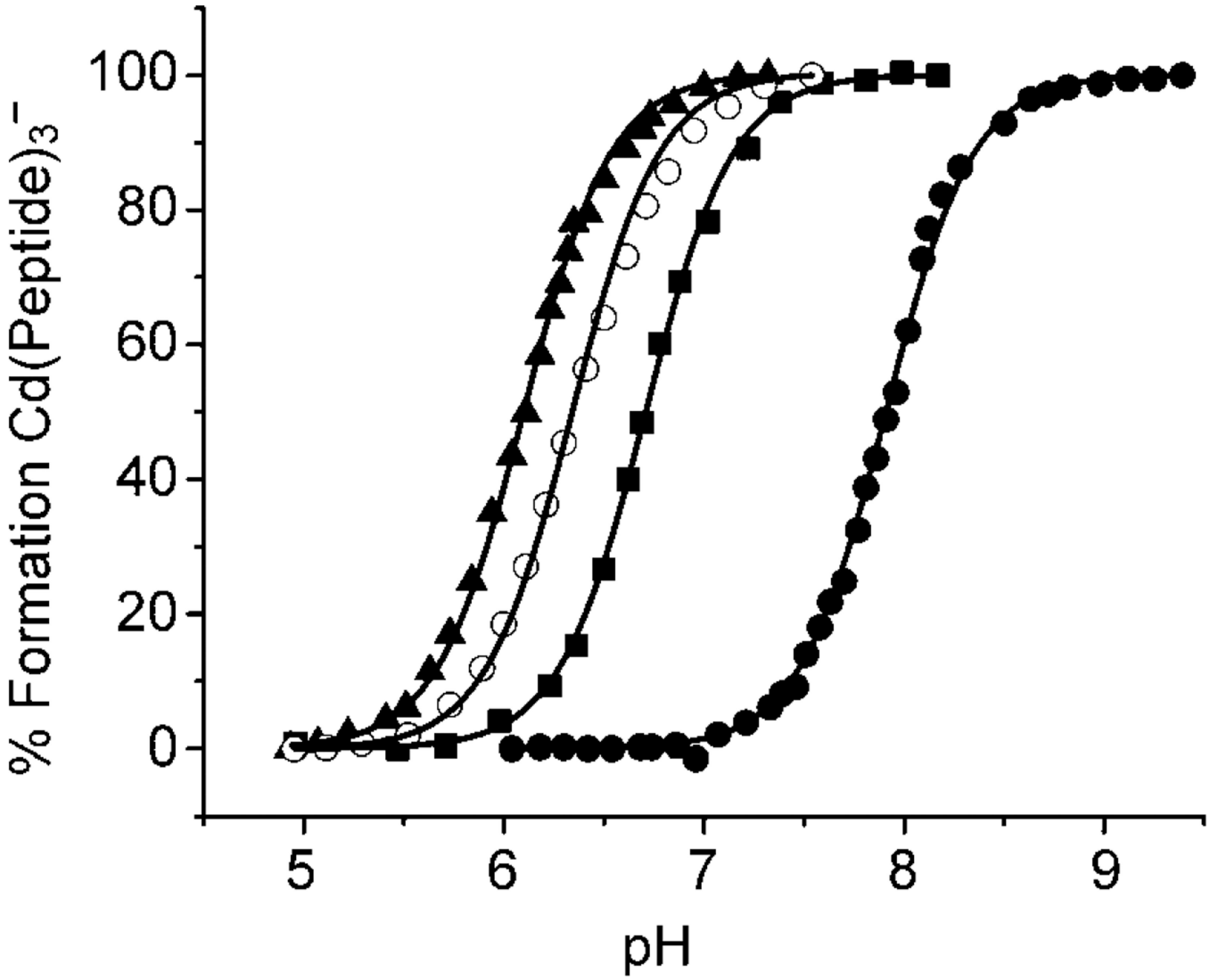

Figure 2.

pH dependence of Cd^{II} binding to **TRI** peptides that contain fully 4-coordinate and fully 3-coordinate binding sites: **TRIL12AL16C** (▲), **TRIL12AL16Pen** (○), **TRIL16C** (■) and **TRIL16Pen** (●). The solid lines represent the fit to the simultaneous deprotonation model.

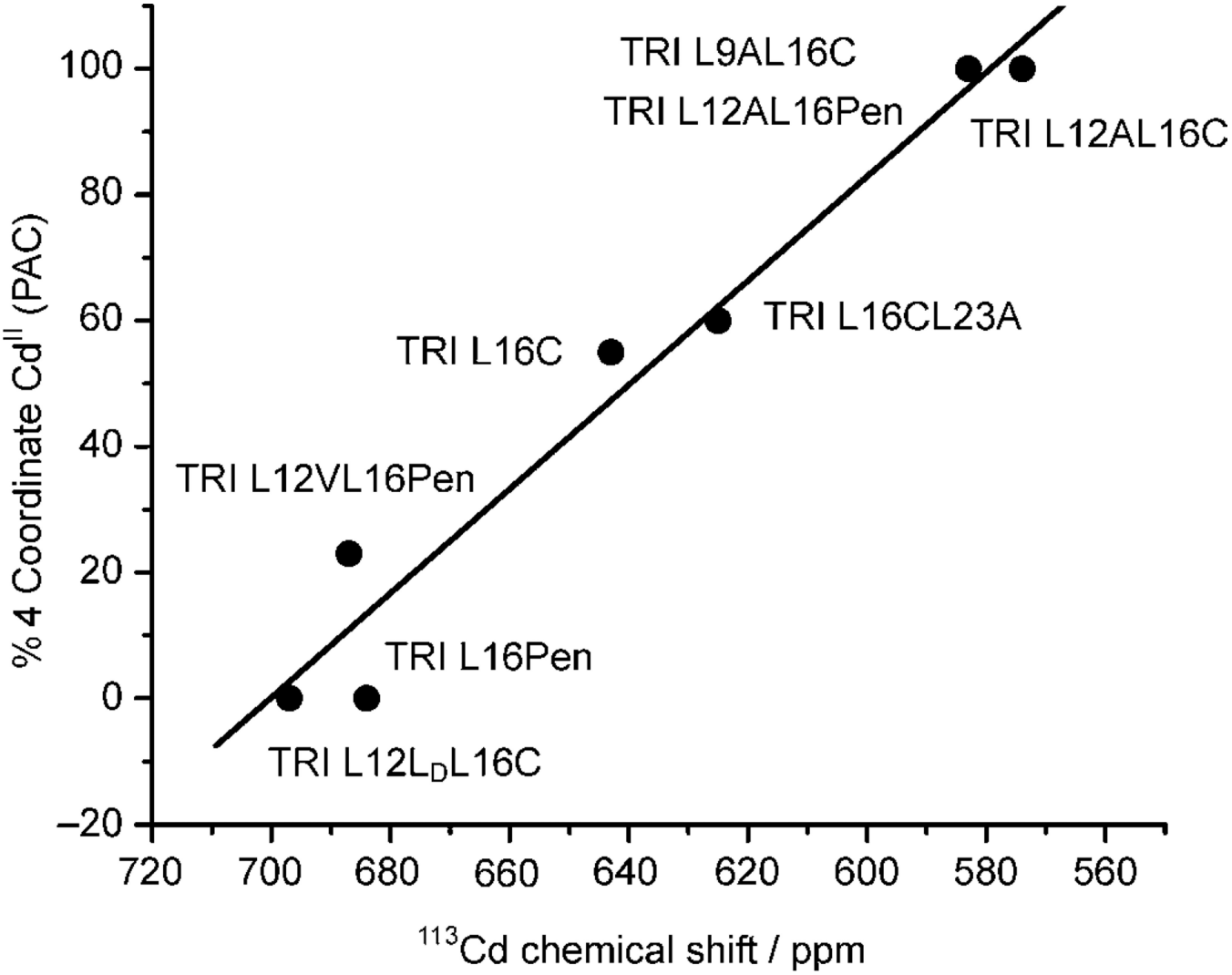

Figure 3.
Plot of the percentage of 4-coordinate Cd^{II} calculated based on the ^{111m}Cd PAC spectroscopic data vs the ^{113}Cd chemical shift for different TRI peptides. The solid line represents the linear fit of the experimental data (r^2 = 0.97414).

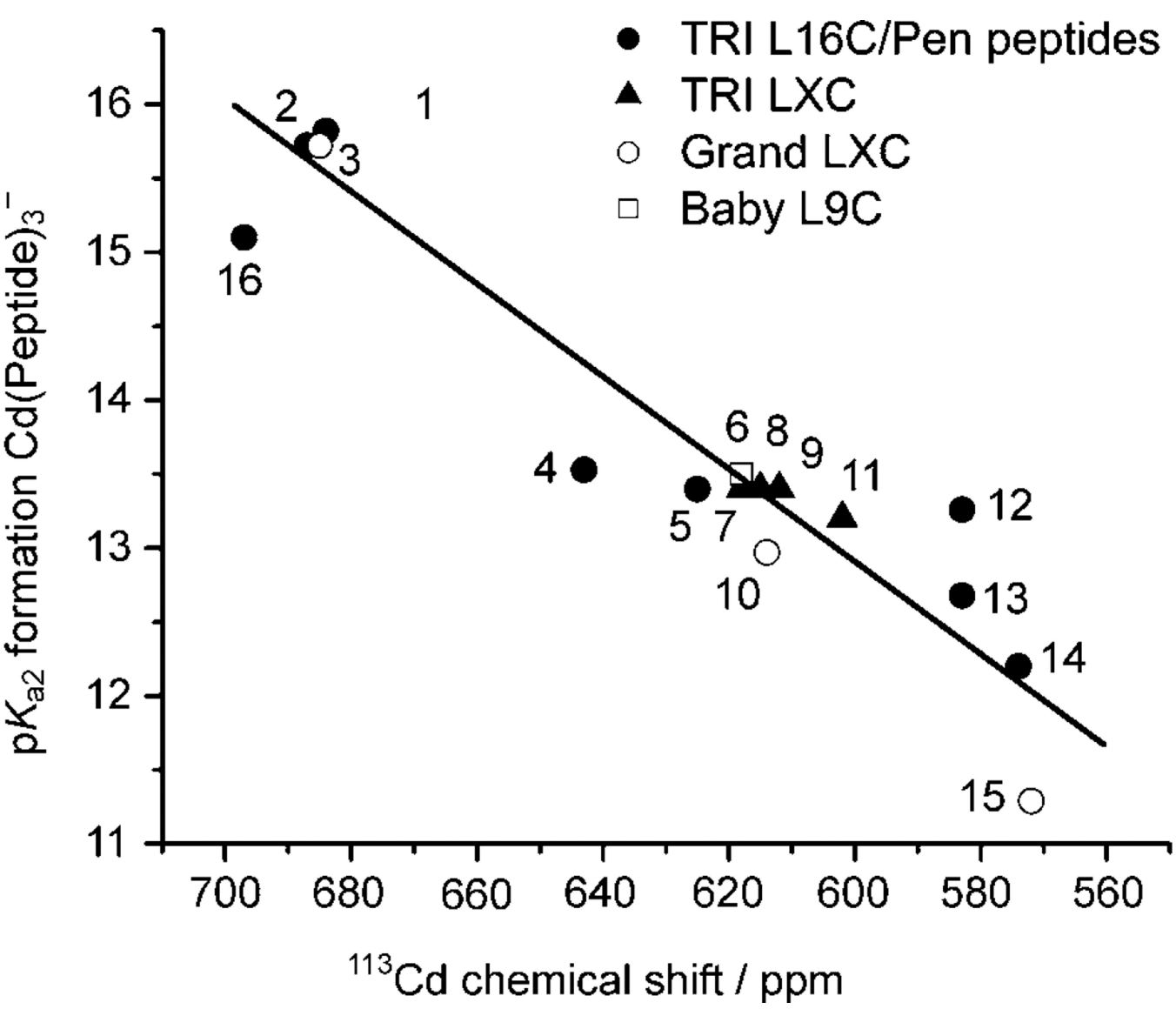

Figure 4.

Plot of the pK_{a2} value for the formation of the complex Cd-(peptide)$_3^-$ vs the ^{113}Cd chemical shift for different peptides of the **TRI** family and derivatives: 1) **TRIL16Pen**; 2) **TRIL12VL16Pen**; 3) **GRANDL16Pen**; 4) **TRIL16CL23A**; 5) **TRIL16C**; 6) **BABYL9C**; 7) **TRIL2WL9C**; 8) **TRIL9C**; 9) **TRIL23C**; 10) **GRANDL9C**; 11) **CSL9C**; 12) **TRIL9AL16C**; 13) **TRIL12AL16Pen**; 14) **TRIL12AL16C**; 15) **GRANDL12AL16C**; 16) **TRIL12L$_D$L16C**. The solid line represents the linear fit of the experimental data (r^2 = 0.9468).

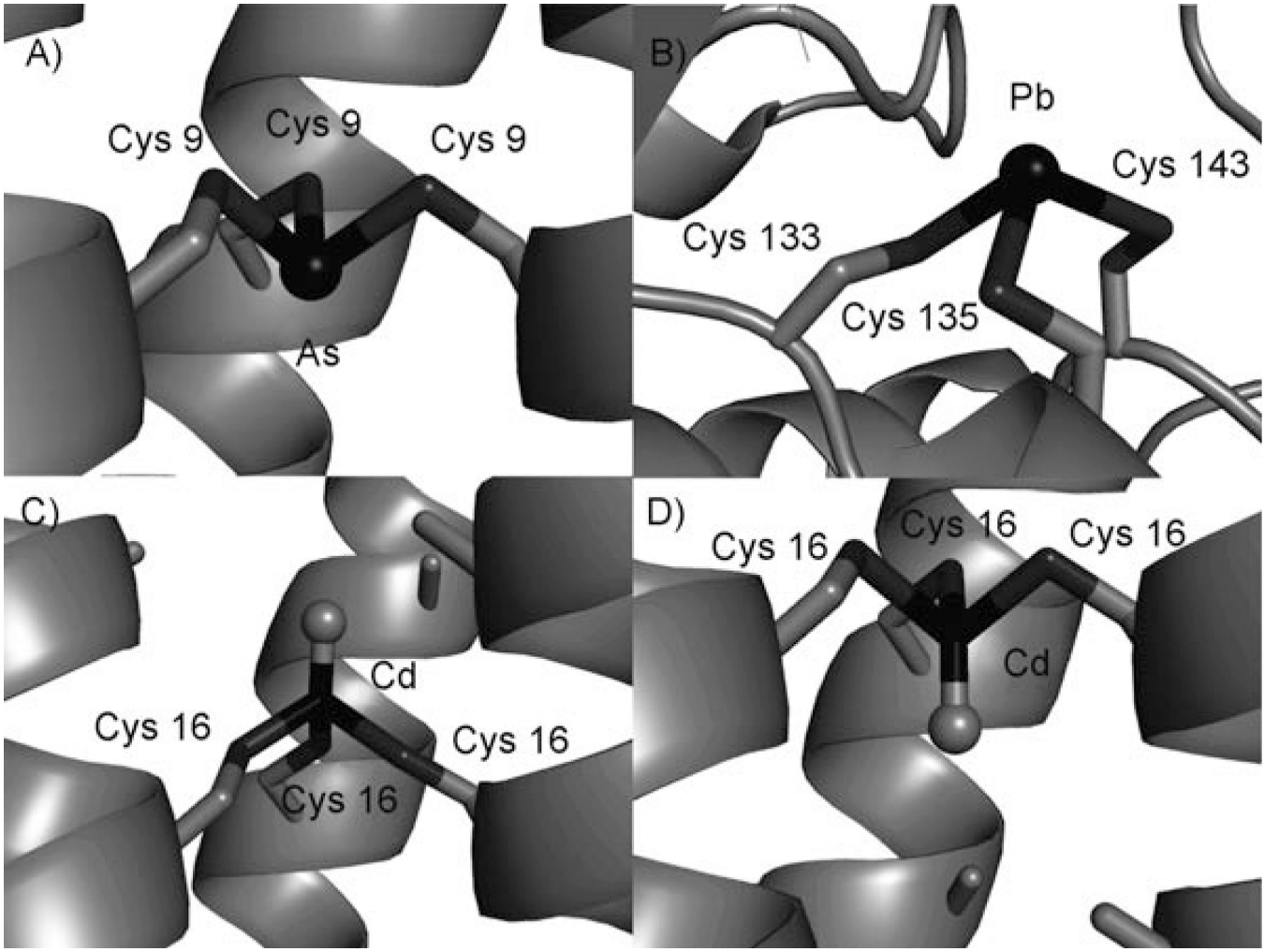

Figure 5.

a) Side-view of As^{III} coordinated to the Cys residues in **CSL9C**.[50] b) Drawing generated using the X-ray crystal structure of Pb^{II} bound to ALAD (pdb: 1QNV).[56] Models, based on the X-ray crystal structure of As(**CSL9C**)$_3$,[50] of CdS_3O bound c) to **TRIL12AL16C** in the *exo* conformation and d) to **TRIL16CL19A** in the *endo* conformation. All figures were generated using PyMOL.

Table 1

Peptide sequence examples of the **TRI** peptide family and derivatives.

Peptide	Sequence[a]
	2 9 12 16 23 30
	abc**d**efg **a**bc**d**efg **a**bc**d**efg **a**bc**d**efg **a**bc**d**efg
TRI	Ac-G LKALEEK LKALEEK LKALEEK LKALEEK G-NH_2
TRIL16C	Ac-G LKALEEK LKALEEK **C**KALEEK LKALEEK G-NH_2
TRIL12AL16C	Ac-G LKALEEK LKA**A**EEK **C**KALEEK LKALEEK G-NH_2
TRIL16Pen	Ac-G LKALEEK LKALEEK **X**KALEEK LKALEEK G-NH_2
TRIL12AL16Pen	Ac-G LKALEEK LKA**A**EEK **X**KALEEK LKALEEK G-NH_2
BABYL9C	Ac-G LKALEEK **C**KALEEK LKALEEK G-NH_2
GRANDL16Pen	Ac-G LKALEEK LKALEEK **X**KALEEK LKAAEEK LKALEEK G-NH_2
CSL9C	Ac-E WEALEKK **C**AALESK LQALEKK LEALEHG-NH_2

[a]**X** = Penicillamine. Residues in bold indicate modifications.

Table 2

Parameters fitted to the PAC data.

Peptide	pH	ω_0 [rad ns^{-1}]	η	$\Delta\omega_0/\omega_0$ (× 100)	$1/\tau_c$ [μs^{-1}]	*A* (×100)	χ_r^2
TRIL16C[a]	8.7	0.337 ± 0.002	0.23 ± 0.02	5.1 ± 0.7	8 ± 5	5.1 ± 0.6	1.10
		0.438 ± 0.004	0.20 ± 0.03	5.4 ± 0.7	3 ± 3	3.6 ± 0.4	
TRIL9AL16C	8.4	0.3332 ± 0.0005	0.19 ± 0.01	2.6 ± 0.2	10.6 ± 0.9	7.5 ± 0.2	1.26
TRIL12AL16C[b]	8.8	0.3405 ± 0.0003	0.141 ± 0.004	1.7 ± 0.1	5.4 ± 0.5	8.3 ± 0.2	1.06
TRIL16CL23A	8.7	0.337 ± 0.002	0.25 ± 0.01	7.6 ± 0.5	4.5 ± 0.5[c]	4.3 ± 0.2	1.20
		0.453 ± 0.002	0.14 ± 0.01	4.1 ± 0.3	4.5 ± 0.5[c]	3.6 ± 0.2	
TRIL16Pen[b]	8.7	0.454 ± 0.001	0.02 ± 0.10	1.6 ± 0.4	19 ± 2	6.9 ± 0.3	0.91
TRIL12VL16Pen	8.7	0.337[d]	0.23[d]	5.1[d]	8[d]	1.5 ± 0.2	0.99
		0.459 ± 0.002	0.20 ± 0.02	2.2 ± 0.8	26 ± 5	5.4 ± 0.5	
TRIL12AL16Pen	8.5	0.339 ± 0.001	0.36 ± 0.01	5.6 ± 0.7	16 ± 2	7.5 ± 0.3	1.11

[a] From ref. [25].

[b] From ref. [30].

[c] The rotational correlation time was constrained to be the same for the two NQIs.

[d] Fixed.

Table 3

Percentage of Cd^{II} species based on ^{111m}Cd PAC spectroscopic data, pK_{a2} values for Cd^{II} binding and ^{113}Cd NMR chemical shifts for the different **TRI** peptides and derivates.

Peptide	% Cd^{II} Species 4-coordinate S_3(O/N)	3-coordinate S_3	Apparent pK_{a2}	^{113}Cd NMR (ppm)
TRIL16C	60	40	13.4 ± 0.2[a]	625[b]
TRIL9AL16C	100	0	13.3 ± 0.2	583[c]
TRIL12AL16C	100	0	12.2 ± 0.2[d]	574[c]
TRIL16CL23A	55	45	13.5 ± 0.2	643
TRIL16Pen	0	100	15.8 ± 0.2[e]	684[f]
TRIL12VL16Pen	22	78	15.7 ± 0.2	687
TRIL12AL16Pen	100	0	12.7 ± 0.2	583
TRIL9C	–	–	13.4 ± 0.2[g]	615[g]
TRIL2WL9C	–	–	13.4 ± 0.2[g]	618[g]
TRIL23C	–	–	13.4 ± 0.2[g]	612[g]
GRANDL9C	–	–	13.0 ± 0.2	614
GRANDL12AL16C	–	–	11.3 ± 0.2	572
GRANDL16Pen	–	–	15.7 ± 0.2[d]	685
BABYL9C	–	–	13.5 ± 0.2	618
CSL9C	–	–	13.2 ± 0.2[g]	602[g]

[a] From ref. [27].

[b] From ref. [25].

[c] From ref. [28].

[d] From ref. [31].

[e] An incorrect value (typo) of 15.3 was reported in ref. [31].

[f] From ref. [30].

[g] From ref. [29].