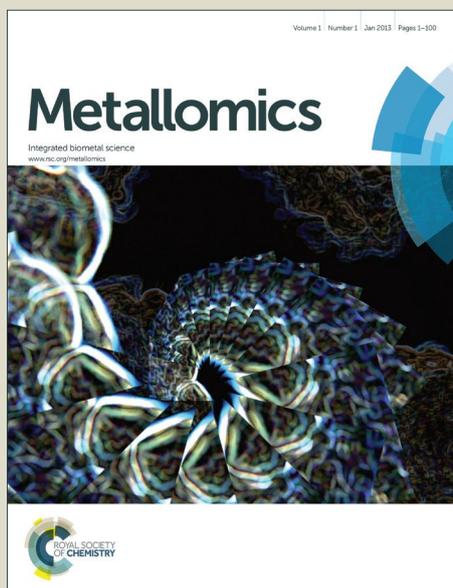


# Metallomics

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

# Variable primary coordination environments of Cd(II) Binding to Three Helix Bundles Provide a Pathway for Rapid Metal Exchange

Alison G. Tebo<sup>a</sup>, Lars Hemmingsen<sup>b</sup>, Vincent L. Pecoraro<sup>c\*</sup>

<sup>a</sup>Program in Chemical Biology, University of Michigan, Ann Arbor, MI 48109

<sup>b</sup>Department of Basic Sciences and Environment, Faculty of Life Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg, Denmark

<sup>c</sup>Department of Chemistry, University of Michigan, Ann Arbor, MI 48109

## Abstract

Members of the ArsR/SmtB family of transcriptional repressors, such as CadC, regulate the intracellular levels of heavy metals like Cd(II), Hg(II), and Pb(II). These metal sensing proteins bind their target metals with high specificity and affinity, however, a lack of structural information about these proteins makes defining the coordination sphere of the target metal difficult. Lingering questions as to the identity of Cd(II) coordination in CadC are addressed via protein design techniques. Two designed peptides with tetrathiolate metal binding sites were prepared and characterized, revealing fast exchange between CdS<sub>3</sub>O and CdS<sub>4</sub> coordination spheres. Correlation of <sup>111m</sup>Cd PAC spectroscopy and <sup>113</sup>Cd NMR spectroscopy suggests that Cd(II) coordinated to CadC is in fast exchange between CdS<sub>3</sub>O and CdS<sub>4</sub> forms, which may provide a mechanism for rapid sensing of heavy metal contaminants by this regulatory protein.

## Introduction

Metal ions are required by all forms of life to sustain cellular processes, but cadmium, lead, and other thiophilic heavy metals can disrupt intracellular processes by binding to protein thiols. DNA-binding metalloregulatory proteins, such as CadC, which is a member of the ArsR/SmtB family of transcriptional repressors, are responsible for sensing the presence of these metals.<sup>1</sup> Highly specific metal-protein interactions govern proper metal homeostasis by the specific uptake, efflux, storage, and trafficking of the metal of interest. While some metals are toxic at nearly all concentrations, others that are often required for life only become toxic at high concentrations. Understanding how these metal-protein interactions can result in detrimental pathologies is important to understanding processes ranging from bioremediation to metal toxicity to neurodegenerative conditions in which an imbalance of metals has been implicated, such as Alzheimer's disease<sup>2</sup>. In particular, the levels of metals are controlled by a complex system of metalloregulatory proteins and chaperones that form specific metal-protein interactions.<sup>1,3</sup> Metalloregulatory proteins often bind only a small subset of metals with high

1  
2 selectivity but how this selectivity is achieved is still not well understood. In general, distorted  
3 or unusual coordination spheres are a common feature of metal sensors.<sup>1,4,5</sup> In particular, CadC  
4 from *S. aureus* pI258 confers resistance to Cd(II), Pb(II), Zn(II), and Bi(III). CadC binds Cd(II)  
5 tetrahedrally and Pb(II) trigonally at the  $\alpha$ 3N (or type 1) site that consists of four cysteines  
6 donated from  $\alpha$ -helices in the core of the protein and from the N terminal loop.<sup>6-8</sup> Mutational  
7 analysis studies have revealed that the cysteines that coordinate Cd(II) and Pb(II) in CadC are  
8 nonequivalent; removal of a particular cysteine residue can still result in metal binding, but the  
9 presence of the bound metal is not “sensed” and the repressor does not dissociate from DNA.<sup>6</sup>  
10 UV-visible spectroscopy and X-Ray Absorption Spectroscopy suggest that Cd(II) coordinates  
11 tetrahedrally with a CdS<sub>4</sub> primary coordination; however, <sup>113</sup>Cd NMR revealed a chemical shift  
12 of 622 ppm,<sup>4</sup> which is significantly lower than that of bonafide Cd-substituted S<sub>4</sub> proteins such  
13 zinc fingers, the structural site metal site in horse liver alcohol dehydrogenase<sup>9</sup> and  
14 rubredoxins<sup>10</sup>. In light of the functional asymmetry of the site due to the non-equivalence of  
15 coordinating cysteines and the fact that no high resolution structures of Cd(II)-bound CadC yet  
16 exist, it has been suggested that Cd(II) bound to CadC assumes a highly distorted tetrahedral  
17 geometry with one very long Cd-S bond,<sup>4,6</sup> but alternatively, the spectroscopic characteristics  
18 could imply that Cd(II) experiences exchange between CdS<sub>3</sub>O and CdS<sub>4</sub> coordination modes.  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31

32 Since few members of the ArsR/SmtB family have been extensively characterized  
33 structurally questions about unusual, intermediate, or distorted coordination spheres are  
34 unresolved. Therefore, our group has turned to protein design methods and specialized  
35 spectroscopic techniques to understand and define the relevant coordination chemistry for toxic  
36 heavy metals. Previous studies have focused on generating thiolate-rich metal binding sites in  
37 three-stranded  $\alpha$ -helical coiled-coils (3SCCs). Amphipathic  $\alpha$ -helices are generated using the  
38 heptad repeat strategy with sequences of seven amino acids (with residues **a-g**) in which the **a**  
39 and **d** positions are occupied by hydrophobic residues that face towards the interior of the  
40 peptide bundle and drive the folding and association of the peptides. Replacement of these  
41 interior-facing hydrophobic residues with metal chelating residues creates the metal binding  
42 site. As(III) coordinated to tris-cysteine sites,<sup>11</sup> as well as trigonal Hg(II) sites<sup>12,13</sup> have been  
43 characterized with these models. Our group has also been able to control the coordination  
44 sphere—both geometry and coordination number—for Cd(II) and Hg(II) within the  
45 hydrophobic interior of peptides, as well as bind Cd(II) in a site-selective manner based on ion  
46 recognition.<sup>14-16</sup>  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2 In particular, the work on Cd(II) has resulted in detailed correlations between Cd(II)  
3 spectroscopic characteristics and the relative proportion of CdS<sub>3</sub> and CdS<sub>3</sub>O coordination  
4 present in a protein,<sup>17</sup> as well as studies on Cd(II) exchange in 3SCCs<sup>18</sup>. More recently, we have  
5 sought to move this chemistry into a single-stranded scaffold, which adopts more of a globular  
6 fold like that of a native protein. A single-stranded construct is capable of enforcing asymmetry,  
7 which is essential to understanding how most native proteins function. Many of the suspected  
8 Cd(II)-binding sensors utilize 3-4 cysteines to generate CdS<sub>3</sub>, CdS<sub>3</sub>O, or distorted CdS<sub>4</sub> sites.<sup>4-6</sup>  
9 To this end,  $\alpha_3$ DIV, a three-helix bundle peptide, was created that can bind Cd(II), Pb(II), and  
10 Hg(II).<sup>19</sup> This construct incorporates three cysteine residues without greatly perturbing the  
11 folding of the peptide, which then are available for metal binding.<sup>19,20</sup> Furthermore, the heptad  
12 repeat structure used in designed  $\alpha$ -helices places potential substitutions either three or four  
13 residues away from the next hydrophobic residue, creating the motifs, CXXL or CXXXL.  
14 Substitution in the  $\alpha_3$ D scaffold, which is inherently asymmetric, could then easily lead to a  
15 controlled number CXXC (or potentially CXXXC) motifs as are found in most native Cd(II)-  
16 binding proteins<sup>1</sup>.

17  
18 Herein, we report the introduction of a fourth cysteine residue in two different locations  
19 in  $\alpha_3$ DIV (Figure 1) to create sites that can mimic native Cd(II)-binding proteins, and in  
20 particular, CadC. Constructing such a site in  $\alpha_3$ DIV presents a significant challenge because  
21 native sites that feature tetrathiolate coordination typically utilize at least one loop or hairpin  
22 secondary structural element to coordinate the metal. Two different peptides with tetrathiolate  
23 metal binding sites were designed and characterized and reveal different proportions of CdS<sub>4</sub>  
24 and CdS<sub>3</sub>O coordination environments. Correlation of UV-visible, <sup>113</sup>Cd NMR, and <sup>111m</sup>Cd  
25 perturbed angular correlation (PAC) spectroscopy suggests that Cd(II) coordinated to CadC is  
26 in fast exchange between CdS<sub>4</sub> and CdS<sub>3</sub>O forms. These observations are used to develop a  
27 model for dynamic exchange that is necessary for proper metalloregulatory protein function.

## 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 Materials and Methods

### 50 51 *Protein Production and Purification*

52  
53  
54 The gene for  $\alpha_3$ DIV-H72C was ordered from Celtek genes (Franklin, TN) subcloned into vector pET15b  
55 with Ap<sup>R</sup> as a selective marker. The construct  $\alpha_3$ DIV-L21C was generated by sequential site directed  
56 mutagenesis of  $\alpha_3$ DIV-H72C by QuikChange kit (Stratagene) to replace the histidine in the 72nd position  
57 (C72H) and incorporating the L21C change. The plasmids were transformed into *E. coli* BL21 (DE3) and plated  
58  
59  
60

1  
2 on LB Amp plates. A single colony was grown overnight in a starter colony and used to inoculate auto-induction  
3 media.<sup>21</sup> The cells were grown for 15-20 hours at 25 °C. The cells were resuspended in 1X PBS with lysozyme  
4 and 5 mM DTT then lysed by microfluidizer. After 30 min of heat denaturation at 55°C, the cell lysate was  
5 acidified and the soluble fraction collected after ultracentrifugation at 15K rpm at 4°C. The pure proteins were  
6 isolated by HPLC on a C18 reverse phase column and lyophilized to yield pure, white powder yielding  
7 approximately 8 mg/L of  $\alpha_3$ DIV-H72C and 20 mg/L of  $\alpha_3$ DIV-L21C. Identity of the proteins was confirmed by  
8 ESI-MS.  
9

#### 10 *UV-visible Spectroscopy*

11  
12 All experiments were carried out using a Cary 100 spectrophotometer. Purified, lyophilized proteins  
13 were resuspended and the concentration determined using the calculated molar extinction coefficient based on  
14 the aromatic residue content. Cadmium titrations were done in 50 mM CHES buffer at pH 8.6 at a peptide  
15 concentration of 20  $\mu$ M. A stock solution of 0.0146 M CdCl<sub>2</sub>, which was standardized by ICP, was titrated into  
16 the peptide solution anaerobically.  
17

#### 18 *<sup>113</sup>Cd Nuclear Magnetic Resonance Spectroscopy*

19  
20 Samples were prepared anaerobically in 10% D<sub>2</sub>O with 0.8 equivalents of enriched <sup>113</sup>Cd with peptide  
21 concentrations of 2-3 mM, and the pH was adjusted to 8.6. Samples were added to a Shigemi solvent-matched  
22 NMR tube and sealed with parafilm. The spectra were collected on a 500 MHz Varian spectrometer using a 90°  
23 pulse (5  $\mu$ s) and 1s acquisition with no delay. Spectra were processed on Mestre-C software with 100 Hz line  
24 broadening.  
25

#### 26 *<sup>111m</sup>Cd Perturbed Angular Correlation Spectroscopy*

27  
28 All perturbed angular correlation (PAC) experiments were performed using a six detector instrument. The  
29 sample was kept at a temperature of 1 °C controlled by a Peltier element. The radioactive cadmium was  
30 produced on the day of the experiment at the University Hospital cyclotron in Copenhagen and extracted as  
31 described previously,<sup>17</sup> except that the HPLC separation of zinc and cadmium was omitted in order to avoid  
32 chloride contamination of the sample. This procedure may lead to zinc contamination of the sample, but the level  
33 of contamination (a few micromolar) should not interfere with the experiment. The <sup>111m</sup>Cd dissolved in ion  
34 exchanged water (10-40  $\mu$ L) was mixed with nonradioactive cadmium acetate and TRIS buffer in appropriate  
35 amounts to achieve the desired final concentrations. The  $\alpha_3$ DIV-H72C or  $\alpha_3$ DIV-L21C peptide was then added  
36 (dissolved in ion-exchanged water), and the sample was left to equilibrate for 10 min to allow for metal binding.  
37 Finally, sucrose was added to produce a 55% w/w solution, and the pH of the solution was adjusted with H<sub>2</sub>SO<sub>4</sub>  
38 or KOH. To measure the pH, a small volume of sample was removed from the solution to avoid chloride  
39 contamination of the sample. The pH was measured at room temperature the following day in the actual sample.  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2 Because of the pH dependence on the temperature of TRIS solutions, the pH of the solution at 1 °C was  
3 calculated using  $\text{pH}(1\text{ °C}) = 0.964[\text{pH}(25\text{ °C})] + 0.86$ . The samples were either used immediately after  
4 preparation or left on ice for a few hours until the measurement was started. All buffers were purged with Ar and  
5 treated so as to lower metal contamination. Time resolution of the measurement was 0.860 ns, and the time pre  
6 channel was 0.562 ns. All fits were carried out with 300 data points, disregarding the 5 first points due to  
7 systematic errors in these. The actual data-analysis is performed on the time dependence of the gamma-gamma  
8 angular correlation function.<sup>22</sup> Two NQIs were included in the analysis. For the minor species the linewidth  
9 ( $\Delta\omega_0/\omega_0$ ) and the asymmetry parameter ( $\eta$ ) were fixed at the value obtained from the other spectrum (where it is  
10 the major species).

## 17 Results

18  
19  
20 There were two likely options for designing a tetrathiolate metal site based on the three metal-  
21 binding cysteine residues from  $\alpha_3\text{DIV}$  as shown in Figure 1.<sup>20</sup> One involved replacing a histidine in a  
22 loop region that previously had been implicated in binding Cd(II),<sup>19</sup> and the other incorporated a  
23 cysteine two residues away from one of  $\alpha_3\text{DIV}$  cysteine residues to generate a single CXXC binding  
24 motif (Table 1). Both peptides fold into a three-helix bundle in the absence of metal (Figure S1). Cd(II)  
25 binds extremely tightly to both  $\alpha_3\text{DIV-H72C}$  and  $\alpha_3\text{DIV-L21C}$  to form 1:1 complexes (Figure S2). The  
26 UV-visible spectra are characterized by ligand-to-metal charge transfer (LMCT) transitions in the UV  
27 region. The complex formed by Cd(II) binding to  $\alpha_3\text{DIV-H72C}$  is characterized by a single broad peak  
28 at 232 nm with a molar extinction coefficient of 25,000  $\text{M}^{-1}\text{cm}^{-1}$  (Figure 2). That of  $\alpha_3\text{DIV-L21C}$   
29 shows a strong transition at 224 nm with a shoulder at 243 nm and molar extinction coefficients of  
30 33,000  $\text{M}^{-1}\text{cm}^{-1}$  and 18,500  $\text{M}^{-1}\text{cm}^{-1}$ , respectively (Figure 2, Table 2).

31  
32  
33 The coordination of Cd(II) in  $\alpha_3\text{DIV-H72C}$  and  $\alpha_3\text{DIV-L21C}$  was assessed using  $^{113}\text{Cd}$  NMR.  
34 Spectra from  $\alpha_3\text{DIV-H72C}$  with 0.8 eq  $^{113}\text{Cd(II)}$  in 10%  $\text{D}_2\text{O}$  show a single peak at 595 ppm, which  
35 can also be seen in Cd(II)-bound  $\alpha_3\text{DIV}$  (Figure 3, Table 2). While very few native proteins exhibit  
36 chemical shifts in this region, previous studies of designed proteins have shown that this chemical shift  
37 range is often associated with  $\text{CdS}_3\text{O}$  coordination.<sup>13,17,23</sup> On the other hand, the chemical shift of  
38  $\alpha_3\text{DIV-L21C}$  is 685 ppm (Figure 2, Table 2), a chemical shift has been more typical of  $\text{CdS}_3$  sites in  
39 designed proteins, although native proteins with pure  $\text{S}_3$  coordination of Cd(II) are not known. The  
40 chemical shift of  $^{113}\text{Cd}$ -substituted rubredoxin has been reported at 730 ppm.<sup>24</sup> Thus, this NMR signal  
41 is most consistent with a high proportion of  $\text{CdS}_4$  coordination that may be undergoing exchange that is  
42 faster than the millisecond timescale of NMR.

1  
2 Since  $^{111m}\text{Cd}$  PAC can resolve multiple species on a faster timescale than NMR, we turned to  
3 this technique to understand the Cd coordination in these two peptides. The data sets for  $\alpha_3\text{DIV-H72C}$   
4 and  $\alpha_3\text{DIV-L21C}$  were initially analyzed with one NQI, capturing the major species, but this did not  
5 give satisfactory fits, and thus two NQIs ( $\omega_0 \sim 0.075$  rad/ns and  $\sim 0.35$  rad/ns) were included for each  
6 sample (Figure 4, Table 3). The  $\omega_0 \sim 0.075$  rad/ns NQI was the major species for  $\alpha_3\text{DIV-H72C}$ , while  
7  $\omega_0 \sim 0.35$  rad/ns was the major species recorded for  $\alpha_3\text{DIV-L21C}$ . This resulted in acceptable fits, and  
8 agreed with visual inspection of the spectra (Figure 4). In other words, for  $\alpha_3\text{DIV-H72C}$  the major  
9 species ( $\omega_0 = 0.356$  rad/ns, 79%) is accompanied by a minor species ( $\omega_0 = 0.082$  rad/ns, 21%), which is  
10 highly similar to the major species recorded (72.5%) for  $\alpha_3\text{DIV-L21C}$ . Conversely, the minor species  
11 (27.5%) recorded for  $\alpha_3\text{DIV-L21C}$  resembles the major species recorded for  $\alpha_3\text{DIV-H72C}$ . The relative  
12 populations for each site were derived from the amplitudes (A) in Table 3. The NQI with  $\omega_0$  around  
13 0.350 rad/ns is comparable to that reported previously in the literature reflecting a  $\text{CdS}_3\text{O}$  coordination  
14 geometry, although with slightly higher frequency.<sup>13,14</sup> The NQI with  $\omega_0$  of about 0.080 rad/ns is highly  
15 similar to that reported in the literature for the structural (Cys<sub>4</sub>) site of horse liver alcohol  
16 dehydrogenase<sup>9</sup> and for a peptide that reproduces this site<sup>25</sup>. Thus, the dominating species for  $\alpha_3\text{DIV-}$   
17 H72C is most likely a  $\text{CdS}_3\text{O}$  coordination geometry, accompanied by a minor fraction, which is most  
18 likely a  $\text{CdS}_4$  coordination geometry. Conversely, the dominating species for  $\alpha_3\text{DIV-L21C}$  is most  
19 likely a  $\text{CdS}_4$  site accompanied by a minor fraction, which is a  $\text{CdS}_3\text{O}$  site.

20  
21 The amplitudes of each species from the fitting of the PAC data, can be used to analyze the  
22 NMR data in order to generate a correlation between the proportion of each species to extract the  
23 “pure” chemical shift for each species. By converting the amplitudes to the fraction represented by each  
24 species, a simple system of equations can be written to calculate the chemical shift of the pure species,  
25  
26

$$685 = 0.7246x + 0.2754y$$

$$595 = 0.2125x + 0.7875y$$

27 where  $x$  and  $y$  are the chemical shift for  $\text{CdS}_3\text{O}$  and  $\text{CdS}_4$ , respectively. Solving this system of  
28 equations predicts a chemical shift of 730 ppm for pure  $\text{CdS}_4$  and 560 ppm for pure  $\text{CdS}_3\text{O}$ , both of  
29 which are reasonable values that fall in line with what has previously been observed for these  $\text{CdS}_4$ <sup>24</sup>  
30 and  $\text{CdS}_3\text{O}$ <sup>17</sup> species, respectively. The observed chemical shift for the  $\text{CdS}_3\text{O}$  species is slightly lower  
31 ( $\sim 20$  ppm) than has been observed for both our previous designed peptides, as well as for native  
32 proteins with this center, but compared over the  $\sim 900$  ppm range accessible by  $^{113}\text{Cd}$  NMR, this is  
33 likely due to systematic error in the calculation of the PAC amplitudes. Indeed, the value of 560 ppm is  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2 within error for that previously determined in our lab based on a comparison of S<sub>3</sub> and S<sub>3</sub>O sites.<sup>17</sup>  
3  
4 These calculations suggest that the observed behavior is due to coalescence of the NMR signals caused  
5  
6 by rapid water/thiolate exchange from the cadmium ion.  
7  
8  
9

## 10 Discussion

11  
12  
13 A tris-thiolate metal binding site was previously incorporated into a single-stranded three-helix  
14 bundle,  $\alpha_3D$ ,<sup>26,27</sup> to generate the construct,  $\alpha_3DIV$ ,<sup>19</sup> which is capable of binding Pb(II), Hg(II), and  
15 Cd(II). A minor species with possible CdS<sub>3</sub>N coordination sphere was measured by <sup>111m</sup>Cd PAC  
16 spectroscopy, suggesting that a nearby histidine residue was capable of interacting with bound Cd(II).  
17 In seeking to incorporate a fourth cysteine residue to create a tetrathiolate metal binding site, a logical  
18 first step would be the replacement of this nearby histidine residue, with a cysteine generating  $\alpha_3DIV$ -  
19 H72C. While histidine has a significantly longer extension than cysteine, the location of this residue on  
20 a flexible loop region could allow for insertion of the thiolate sulfur into the Cd(II) first coordination  
21 sphere. As an alternate approach, we also identified another residue, Leu21, which is located two  
22 residues away from Cys18, one of the cysteines incorporated into  $\alpha_3DIV$ . This site would incorporate a  
23 CXXC motif, much as is found in zinc fingers, rubredoxin, and Cd(II)-sensing proteins. These two  
24 constructs represent two reasonable approaches to generating S<sub>4</sub> sites, and characterizing them could  
25 also provide some insights into the basic requirements for tetrahedral metal binding sites.  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36

37 The UV-visible spectroscopy of both  $\alpha_3DIV$ -H72C and  $\alpha_3DIV$ -L21C are consistent with Cd(II)  
38 coordinated to cysteine thiolates. Both of these constructs have much higher molar extinction  
39 coefficients than previously designed Cd(II)-binding peptides from our laboratory,<sup>13,19,23,28</sup> but are  
40 consistent with native Cd(II)-binding proteins<sup>6</sup>. Previous studies have elucidated that molar extinction  
41 coefficients increase linearly with respect to the increasing number of thiolates coordinated to Cd(II),<sup>29</sup>  
42 so the increase in molar extinction coefficient observed for these peptides compared to other designed  
43 peptides (e.g., Cd(**TRIL12AL16C**)<sub>3</sub>)<sup>23</sup> is consistent with increased cysteine coordination. CadC shows  
44 a single, broad peak at 238 nm with a molar extinction coefficient of 25,000 M<sup>-1</sup>cm<sup>-1</sup>.<sup>4</sup> This differs  
45 greatly from the absorption of Cd-substituted rubredoxin, which is centered at 232 nm with three  
46 resolvable transitions at 245 nm, 229 nm, and 213 nm.<sup>29</sup> The molar extinction coefficient of the lowest  
47 energy transition can be used to estimate the number of thiolates bonded to the Cd(II) based on ~6000  
48 M<sup>-1</sup>cm<sup>-1</sup> per thiolate.<sup>29</sup> Other studies, including those on CadC, have used this rule of thumb even when  
49 the spectrum only has a single broad peak and the lowest energy transition was not resolvable.<sup>6</sup> The  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2 extinction coefficient of the shoulder at 243 nm of  $\alpha_3\text{DIV-L21C}$  is lower than the  $24,000 \text{ M}^{-1}\text{cm}^{-1}$   
3 benchmark for tetrathiolate coordination, which could signal that less than 100% of the protein in the  
4 sample binds to Cd(II) with four thiolates. The lack of multiple resolvable transitions in CadC and  
5  $\alpha_3\text{DIV-H72C}$  is interesting and suggests that  $\alpha_3\text{DIV-H72C}$  is in some ways a good spectroscopic model  
6 for CadC. The band structure found in rubredoxins and zinc fingers likely arises from the tetrahedral  
7 geometry of the site, and the lack of this structure in CadC indicates that the Cd(II) coordination could  
8 be highly distorted with one very long Cd—S bond, or in equilibrium between  $\text{CdS}_3\text{O}$  and  $\text{CdS}_4$ ,  
9 creating the broader transition at an intermediate energy (238 nm vs. 229 nm and 245 nm as found in  
10 rubredoxin). This appears to also be the case in  $\alpha_3\text{DIV-H72C}$ , which features a similar broad band at an  
11 intermediate energy (232 nm vs. 224 nm and 243 nm as found in  $\alpha_3\text{DIV-L21C}$ ). In comparison to  
12  $\alpha_3\text{DIV}$ ,  $\alpha_3\text{DIV-H72C}$  has a higher molar extinction coefficient (18,200 vs. 25,000), which supports the  
13 concept that this may represent a mixture of two coordination environments and that some greater  
14 amount of thiolate coordination is present.  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24

25  
26  $^{113}\text{Cd}$  NMR was used to investigate Cd(II) coordination in  $\alpha_3\text{DIV-H72C}$  and  $\alpha_3\text{DIV-L21C}$ . The  
27 chemical shifts of  $\alpha_3\text{DIV-H72C}$  and  $\alpha_3\text{DIV-L21C}$  were 595 ppm and 685 ppm, respectively. As a  
28 result of this data, the original assignments for  $\alpha_3\text{DIV}$   $^{113}\text{Cd}$  NMR have been revisited. Initially it was  
29 thought that the signal at 595 ppm represented a  $\text{CdS}_3\text{N}$  species while the signal at 583 ppm reflected a  
30  $\text{CdS}_3\text{O}$  species. Yet comparison of  $^{113}\text{Cd}$  NMR data from  $\alpha_3\text{DIV-H72C}$  and  $\alpha_3\text{DIV}$  reveals that the 595  
31 ppm resonance persists in the two proteins indicating that this signal represents the  $\text{CdS}_3\text{O}$  species.  
32 While very few native proteins exhibit chemical shifts in this region, studies in our lab have shown that  
33 this chemical shift range is often associated with  $\text{CdS}_3\text{O}$ . A series of studies on rubredoxin mutants  
34 with cysteine to serine mutations measured the chemical shifts of these  $\text{S}_3\text{O}$  sites to be between 605-  
35 645 ppm, while the chemical shift of  $^{113}\text{Cd}$ -substituted rubredoxin has been reported at 730 ppm.<sup>24</sup>  
36 However, there are Cd-substituted  $\text{S}_4$  sites in native proteins that have shown chemical shifts lower  
37 than 700 ppm, although many of these tend to be  $\text{S}_4$  bimetallic bridged structures.<sup>10</sup> The origin of the  
38 chemical shift of 685 ppm for  $\alpha_3\text{DIV-L21C}$  could be due to a distorted site or to an equilibrium  
39 between  $\text{CdS}_3\text{O}$  and  $\text{CdS}_4$  forms that exchange faster than the NMR timescale, leading to coalescence  
40 of the signals from the two species to a single, intermediate chemical shift.  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52

53  
54 The  $^{111\text{m}}\text{Cd}$  PAC data indicate that both  $\text{CdS}_3\text{O}$  and  $\text{CdS}_4$  species are found in both peptides in  
55 inverse proportions. Taken together, the NMR and PAC spectroscopic data suggest that  $\alpha_3\text{DIV-H72C}$   
56 adopts a primarily  $\text{CdS}_3\text{O}$  coordination with a small proportion of  $\text{CdS}_4$ . The UV-visible spectroscopy  
57 supports this model as the spectrum is characterized by a single broad transition. The molar extinction  
58  
59  
60

1  
2 coefficient is higher than those previously reported for these sites by our group, but this could be due to  
3 the small proportion of CdS<sub>4</sub>, which can contribute up to 6000 M<sup>-1</sup>cm<sup>-1</sup> to the molar extinction  
4 coefficient.<sup>29,30</sup> The PAC and UV-visible spectroscopy indicate that the major species for α<sub>3</sub>DIV-L21C  
5 is CdS<sub>4</sub>. The chemical shift for this peptide is smaller than that for a Cd-substituted classically  
6 tetrahedral protein, such as rubredoxin. Given the two species in the PAC data and the good agreement  
7 in energy for the LMCT, it seems most likely that coalescence of the signals causes the smaller  
8 chemical shift in the <sup>113</sup>Cd NMR. Assuming the molar extinction coefficient for the resolved low  
9 energy transition in α<sub>3</sub>DIV-L21C is only due to CdS<sub>4</sub>, a correction can be derived to calculate the real  
10 molar extinction coefficient for the CdS<sub>4</sub> species. This calculation yields ε<sub>243nm</sub> of 25,500 M<sup>-1</sup>cm<sup>-1</sup>,  
11 which reflects the expected molar extinction coefficient for a CdS<sub>4</sub> site.  
12  
13  
14  
15  
16  
17  
18  
19

20  
21 Based on the correlation between the PAC amplitudes and the NMR data, the chemical shift of  
22 the pure CdS<sub>3</sub>O and CdS<sub>4</sub> species was calculated to be 560 ppm and 730 ppm, respectively. These  
23 values agree very well with what has previously been determined for these species. The value for  
24 CdS<sub>3</sub>O coordination is within 20 ppm of what has previously been determined in our lab for a  
25 correlation between CdS<sub>3</sub> and CdS<sub>3</sub>O coordinations,<sup>17</sup> which is within the margin of error for the  
26 determination of the PAC amplitudes. This rationale and calculated values for the pure species can be  
27 applied to CadC. The NMR signal for CadC, 622 ppm, is then made up of two signals at differing  
28 proportions. The equation for this system can be written as,  
29  
30  
31  
32  
33  
34

$$622 = 734 + 557(1 - x)$$

35  
36  
37 where  $x$  is the proportion of S<sub>4</sub> species. This analysis suggests that the NMR signal for CadC is  
38 produced by coalescence between the two forms, where 40% exists as S<sub>4</sub> and 60% exists as S<sub>3</sub>O. This  
39 NMR-derived speciation model also explains the λ<sub>max</sub> and molar extinction coefficient from the UV  
40 spectra of CadC. Finally, the EXAFS data recorded for CadC are fitted almost equally well with a CdS<sub>4</sub>  
41 and a CdS<sub>3</sub>O coordination sphere, but with a very large Debye-Waller factor for the oxygen.<sup>4</sup> Thus, we  
42 conclude that the results from the α<sub>3</sub>DIV-H72C and α<sub>3</sub>DIV-L21C provide compelling evidence  
43 supporting a rapid exchange between S<sub>3</sub>O and S<sub>4</sub> species in CadC, and that this model provides a  
44 coherent interpretation of UV-absorption, <sup>113</sup>Cd NMR, and EXAFS spectroscopic data recorded for  
45 CadC.  
46  
47  
48  
49  
50  
51  
52  
53

54  
55 The function of CadC is to serve as a sensitive metal ion sensor that can respond rapidly to  
56 variations of Cd(II) concentrations in cells. When Cd(II) levels are extremely low, it must repress  
57 synthesis of the heavy metal detoxification pathway by binding to DNA, while falling off the repressor  
58  
59  
60

1 region when concentrations of Cd(II) increase. As this sensor works at picomolar Cd(II)  
2 concentrations, this requires CadC to have a very high affinity for its target metal. However, this switch  
3 must be reversible as when Cd(II) levels are diminished, the protein must be reused to repress gene  
4 expression. Rapid exchange between two coordination spheres can serve as a functional feature of  
5 CadC such that this fast exchange can facilitate transfer of Cd(II) from small molecule ligands, such as  
6 glutathione, to the metal binding site. Furthermore, the dynamic nature of the site can allow it to sense  
7 small changes in the intracellular concentration of Cd(II) resulting in a highly sensitive and responsive  
8 sensor. Similarly, PAC data clearly demonstrate the presence of two co-existing coordination spheres  
9 for Ag(I) and possibly for Cd(II) bound to the ArsR/SmtB family member BxmR.<sup>31</sup> Thus, it seems  
10 plausible that an intrinsic feature of such metalloregulating proteins is the ability to have a dynamic  
11 coordination site that can facilitate exchange so that the protein does not irreversibly lock a metal into  
12 the regulatory domain.  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23

24 In summary, we have reported the design and characterization of two three-helix bundles with  
25 four cysteine residues. Both peptides bind Cd(II) tightly in a 1:1 complex that is characterized by high  
26 energy LMCT. The two peptides,  $\alpha_3$ DIIV-H72C and  $\alpha_3$ DIIV-L21C, were then characterized with <sup>113</sup>Cd  
27 NMR, suggesting that each assumes a different coordination. <sup>111m</sup>Cd PAC revealed that both peptides  
28 coordinate Cd(II) in a mixture of CdS<sub>3</sub>O and CdS<sub>4</sub> with opposite major species such that the major  
29 species of  $\alpha_3$ DIIV-H72C was CdS<sub>3</sub>O and  $\alpha_3$ DIIV-L21C was CdS<sub>4</sub>. Correlating the NMR chemical shifts  
30 and the proportion of each species derived from the PAC fits results in the pure species being  
31 represented by chemical shifts of 730 ppm and 560 ppm for CdS<sub>4</sub> and CdS<sub>3</sub>O, respectively. Applying  
32 these values to CadC suggests that CadC also coordinated Cd(II) as a mixture of CdS<sub>3</sub>O and CdS<sub>4</sub>  
33 under rapid exchange with 60% CdS<sub>3</sub>O and 40% CdS<sub>4</sub>, which agrees well with reported spectroscopy  
34 for this protein. This data suggests a model for metalloregulatory proteins in which apparently distorted  
35 coordination spheres are actually in fast exchange between two different forms, which may function to  
36 facility metal binding from low molecular weight compounds in the cytosol and sense small changes in  
37 the concentration of metals.  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48

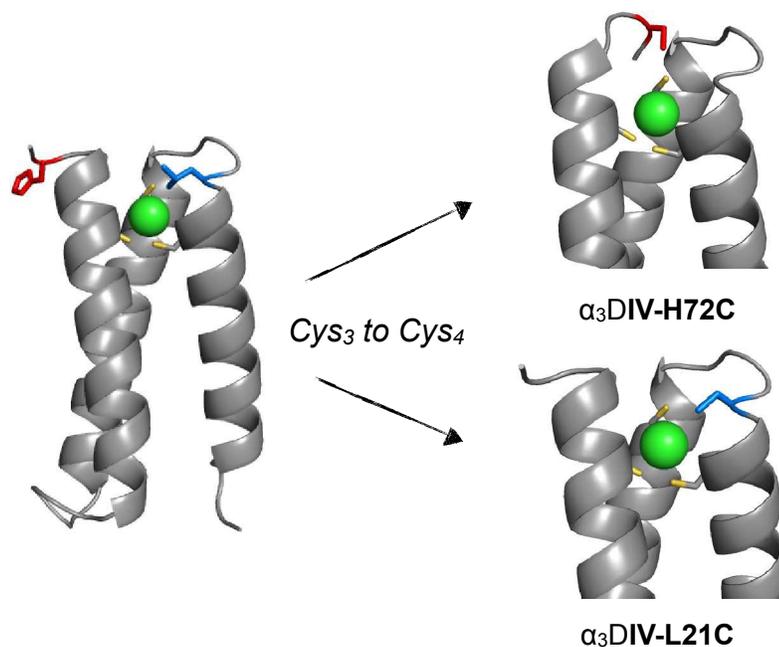
#### 49 References:

- 50 1 Z. Ma, F. E. Jacobsen and D. P. Giedroc, *Chem. Rev.*, 2009, **109**, 4644–4681.
- 51 2 E. Gaggelli, H. Kozłowski, D. Valensin and G. Valensin, *Chem. Rev.*, 2006, **106**, 1995–2044.
- 52 3 J. M. Argüello, D. Raimunda and M. González-Guerrero, *J. Biol. Chem.*, 2012, **287**, 13510–13517.
- 53 4 L. S. Busenlehner, N. J. Cosper, R. A. Scott, B. P. Rosen, M. D. Wong and D. P. Giedroc, *Biochemistry*,  
54 2001, **40**, 4426–4436.
- 55 5 Y. Wang, J. Kendall, J. S. Cavet and D. P. Giedroc, *Biochemistry*, 2010, **49**, 6617–6626.
- 56 6 L. S. Busenlehner, T.-C. Weng, J. E. Penner-Hahn and D. P. Giedroc, *J. Mol. Biol.*, 2002, **319**, 685–701.
- 57  
58  
59  
60

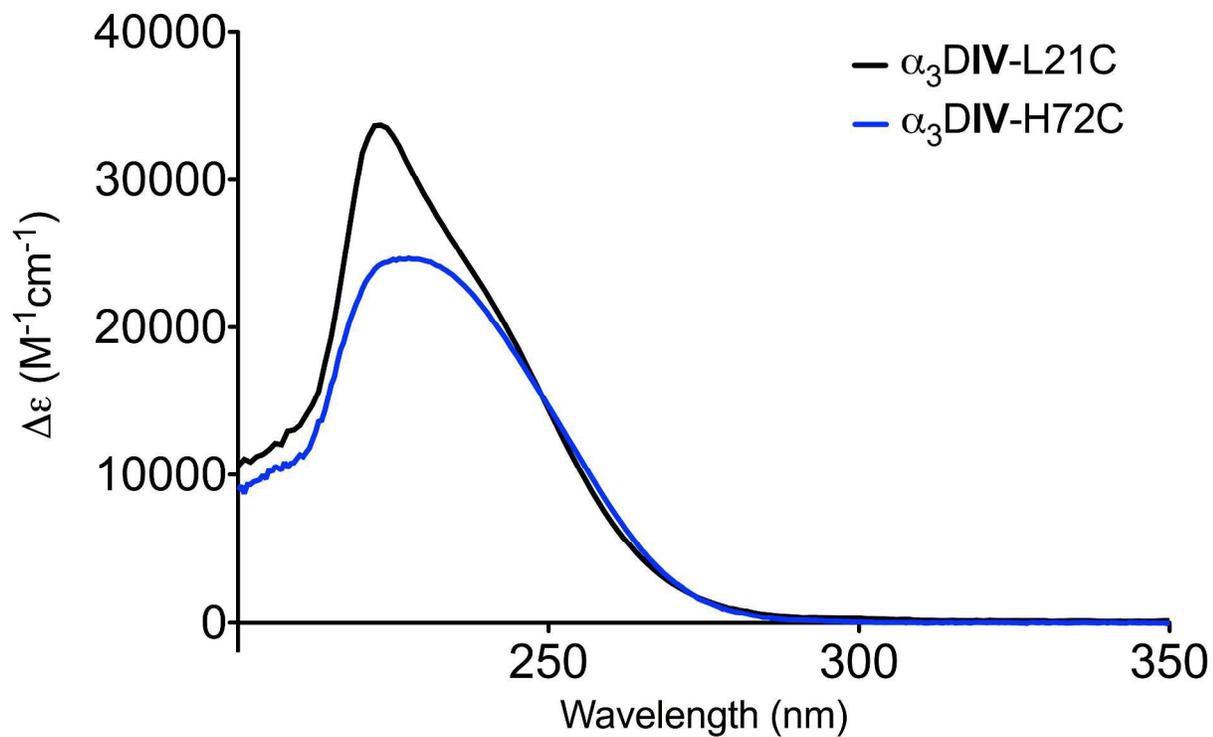
- 1  
2 7 J. Ye, A. Kandegedara, P. Martin and B. P. Rosen, *J. Bacteriol.*, 2005, **187**, 4214–4221.  
3 8 L. S. Busenlehner and D. P. Giedroc, *J. Inorg. Biochem.*, 2006, **100**, 1024–1034.  
4 9 L. Hemmingsen, R. Bauer, M. J. Bjerrum, M. Zeppezauer, H. W. Adolph, G. Formicka and E. Cedergren-  
5 Zeppezauer, *Biochemistry*, 1995, **34**, 7145–7153.  
6 10 I. M. Armitage, T. Drakenberg and B. Reilly, in *Cadmium: From Toxicity to Essentiality*, eds. A. Sigel, H.  
7 Sigel and R. K. Sigel, Springer Netherlands, Dordrecht, 2012, vol. 11, pp. 117–144.  
8 11 D. S. Touw, C. E. Nordman, J. A. Stuckey and V. L. Pecoraro, *Proceedings of the National Academy of*  
9 *Sciences*, 2007, **104**, 11969–11974.  
10 12 G. R. Dieckmann, D. K. McRorie, D. L. Tierney, L. M. Utschig, C. P. Singer, T. V. O'Halloran, J. E.  
11 Penner-Hahn, W. F. DeGrado and V. L. Pecoraro, *J. Am. Chem. Soc.*, 1997, 6195–6196.  
12 13 M. Matzapetakis, B. T. Farrer, T.-C. Weng, L. Hemmingsen, J. E. Penner-Hahn and V. L. Pecoraro, *J. Am.*  
13 *Chem. Soc.*, 2002, **124**, 8042–8054.  
14 14 A. F. A. Peacock, L. Hemmingsen and V. L. Pecoraro, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 16566–  
15 16571.  
16 15 O. Iranzo, C. Cabello and V. L. Pecoraro, *Angew. Chem. Int. Ed. Engl.*, 2007, **46**, 6688–6691.  
17 16 O. Iranzo, S. Chakraborty, L. Hemmingsen and V. L. Pecoraro, *J. Am. Chem. Soc.*, 2011, **133**, 239–251.  
18 17 O. Iranzo, T. Jakusch, K.-H. Lee, L. Hemmingsen and V. L. Pecoraro, *Chem. Eur. J.*, 2009, **15**, 3761–3772.  
19 18 S. Chakraborty, O. Iranzo, E. R. P. Zuiderweg and V. L. Pecoraro, *J. Am. Chem. Soc.*, 2012, **134**, 6191–  
20 6203.  
21 19 S. Chakraborty, J. Y. Kravitz, P. W. Thulstrup, L. Hemmingsen, W. F. DeGrado and V. L. Pecoraro,  
22 *Angew. Chem. Int. Ed. Engl.*, 2011, **50**, 2049–2053.  
23 20 J. S. Plegaria, S. Dzul, E. R. P. Zuiderweg, T. L. Stemmler and V. L. Pecoraro, *Biochemistry*, 2015.  
24 21 F. W. Studier, *Protein Express. Purif.*, 2005, **41**, 207–234.  
25 22 L. Hemmingsen, K. N. Sas and E. Danielsen, *Chem. Rev.*, 2004, **104**, 4027–4062.  
26 23 O. Iranzo, D. Ghosh and V. L. Pecoraro, *Inorg. Chem.*, 2006, **45**, 9959–9973.  
27 24 Z. Xiao, M. J. Lavery, M. Ayhan, S. D. B. Scrofani, M. C. J. Wilce, J. M. Guss, P. A. Tregloan, G. N.  
28 George and A. G. Wedd, *J. Am. Chem. Soc.*, 1998, **120**, 4135–4150.  
29 25 U. Heinz, L. Hemmingsen, M. Kiefer and H.-W. Adolph, *Chem. Eur. J.*, 2009, **15**, 7350–7358.  
30 26 S. T. Walsh, H. Cheng, J. W. Bryson, H. Roder and W. F. DeGrado, *Proc. Natl. Acad. Sci. U. S. A.*, 1999,  
31 **96**, 5486–5491.  
32 27 J. W. Bryson, J. R. Desjarlais, T. M. Handel and W. F. DeGrado, *Protein Sci.*, 1998, **7**, 1404–1414.  
33 28 M. Matzapetakis, D. Ghosh, T.-C. Weng, J. E. Penner-Hahn and V. L. Pecoraro, *J. Biol. Inorg. Chem.*,  
34 2006, **11**, 876–890.  
35 29 C. J. Henehan, D. L. Pountney, M. Vasák and O. Zerbe, *Protein Sci.*, 1993, **2**, 1756–1764.  
36 30 D. L. Pountney, R. P. Tiwari and J. B. Egan, *Protein Sci.*, 1997, **6**, 892–902.  
37 31 T. Liu, X. Chen, Z. Ma, J. Shokes, L. Hemmingsen, R. A. Scott and D. P. Giedroc, *Biochemistry*, 2008, **47**,  
38 10564–10575.  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

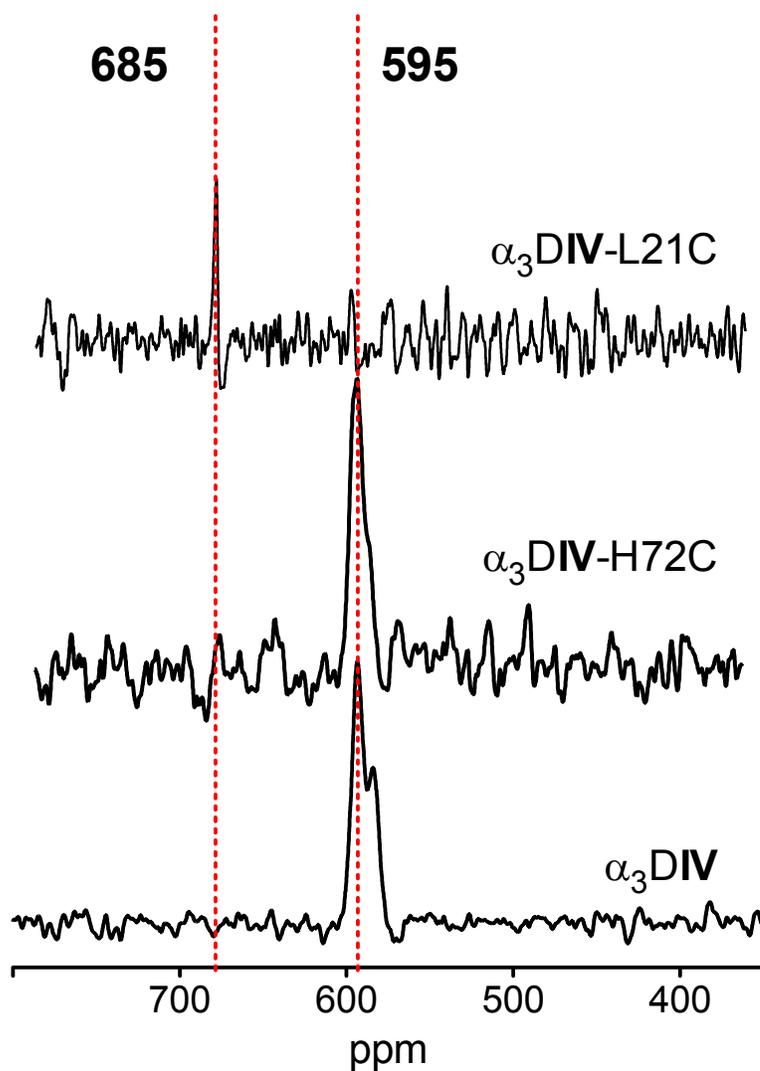
Figures:



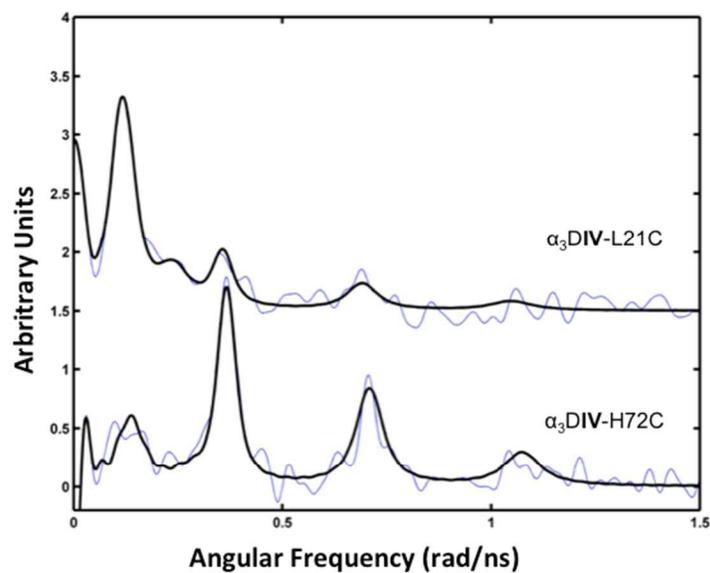
**Figure 1** Design strategy for the generation of a tetrathiolate metal-binding motif in the  $\alpha_3\text{DIV}$  scaffold. Here, the substitution of His72 (red residue), which coordinates Cd(II) in  $\alpha_3\text{DIV}$  could allow for tetrathiolate coordination of Cd(II). An alternative approach substitutes Leu21 (pictured in blue) to create a CXXC binding motif. Models are based on PDB: 2MTQ



**Figure 2** UV-visible spectroscopy of Cd(II) bound to  $\alpha_3$ DIV-H72C and  $\alpha_3$ DIV-L21C in 50 mM CHES buffer at pH 8.6.



**Figure 3**  $^{113}\text{Cd}$  NMR reveals different coordination environments for  $\alpha_3\text{DIV-H72C}$  and  $\alpha_3\text{DIV-L21C}$ , while the coordination environments in  $\alpha_3\text{DIV}$  and  $\alpha_3\text{DIV-H72C}$  are similar. All experiments were conducted with 2-3 mM peptide and 0.8 eq. of enriched  $^{113}\text{Cd}$  in 10%  $\text{D}_2\text{O}$ .



**Figure 4**  $^{111m}\text{Cd}$  PAC spectra for  $\alpha_3\text{DIV-L21C}$  (top) and  $\alpha_3\text{DIV-H72C}$  (bottom) with the data (fine line) and fit (thick line). Experiment was performed at 300  $\mu\text{M}$  peptide, 20 mM TRIS and a peptide/Cd(II) ratio of 1:12.

**Table 1**

Peptide	Sequence
$\alpha_3$ DIV	MGS WAEFKQR LAAIKTR CQAL GG SEAECAAFEKE IAAFESE LQAY KGKG NPE VEALRKE AAAIRDE CQAY RHN
$\alpha_3$ DIV- H72C	MGS WAEFKQR LAAIKTR LQAC GG SEAECAAFEKE IAAFESE LQAY KGKG NPE VEALRKE AAAIRDE CQAY RCN GSGC
$\alpha_3$ DIV- L21C	MGS WAEFKQR LAAIKTR CQAC GG SEAECAAFEKE IAAFESE LQAY KGKG NPE VEALRKE AAAIRDE CQAY RHN GSGC

**Table 2** Tabulated values for UV-visible spectroscopy and  $^{113}\text{Cd}$  NMR

Peptide	UV-vis [nm ( $\epsilon$ )]	$^{113}\text{Cd}$ NMR [ppm]
$\alpha_3\text{DIV}^{\text{a}}$	232 (18,200)	583, 595
$\alpha_3\text{DIV-H72C}$	232 (25,000)	595
$\alpha_3\text{DIV-L21C}$	224 (33,000)	685
	243 (18,500)	
$\text{CadC}^{\text{b}}$	238 (25,000)	620
Rubredoxin	213 (25,000) 229 (45,000) 245 (26,000) <sup>c,d</sup>	734 <sup>e</sup>
TRIL12AL16C <sup>f</sup>	231 (21,200)	574
TRIL16C <sup>g</sup>	232 (22,600)	625

<sup>a</sup> ref 19 <sup>b</sup> ref 6 <sup>c</sup> ref 29 <sup>d</sup> values at 213 nm and 229 nm estimated from visual inspection of published spectra <sup>e</sup> ref 24 <sup>f</sup> ref 23 <sup>g</sup> ref 28

**Table 3** Parameters fitted to PAC-data.

Peptide	pH	$\omega_0$ (rad/ns)	$\eta$	$\Delta\omega_0/\omega_0$ $\times 100$	$1/\tau_c$ $\mu\text{s}^{-1}$	A $\times 100$	$\chi_r^2$
$\alpha_3\text{DIV-H72C}^{\text{a}}$	8.6	$0.082 \pm 0.003$	$0.79^{\text{f}}$	15 <sup>f</sup>	$4.3 \pm 0.7$	$1.7 \pm 0.2$	1.26
		$0.356 \pm 0.001$	$0.16 \pm 0.01$	$4.2 \pm 0.5$	$4.3 \pm 0.7$	$6.3 \pm 0.4$	
$\alpha_3\text{DIV-L21C}^{\text{a}}$	8.6	$0.071 \pm 0.002$	$0.79 \pm 0.06$	$15 \pm 3$	$5.2 \pm 0.8$	$5.0 \pm 0.2$	1.28
		$0.347 \pm 0.003$	$0.16^{\text{f}}$	4 <sup>f</sup>	$5.2 \pm 0.8$	$1.9 \pm 0.2$	
<b>TRIL16C</b> <sup>b, c</sup>	8.7	$0.337 \pm 0.002$	$0.23 \pm 0.02$	$5.1 \pm 0.7$	$8 \pm 5$	$5.1 \pm 0.6$	1.10
		$0.438 \pm 0.004$	$0.20 \pm 0.03$	$5.4 \pm 0.7$	$3 \pm 3$	$3.6 \pm 0.4$	
<b>TRIL12AL16C</b> <sup>d, c</sup>	8.8	$0.3405 \pm 0.0003$	$0.141 \pm 0.004$	$1.7 \pm 0.1$	$5.4 \pm 0.5$	$8.3 \pm 0.2$	1.06
$\text{C}_4$ peptide <sup>f</sup>		$0.075 \pm 0.001$	$0.60 \pm 0.09$	$16 \pm 5$	$38 \pm 5$		
LADH $\text{C}_4$ site <sup>g</sup>		$0.065 \pm 0.001$	$0.81 \pm 0.06$	$10 \pm 3$	$8.7 \pm 0.1$		

f: fixed. <sup>a</sup> this work; <sup>b</sup> from ref 17 <sup>c</sup> this peptide is 60% CdS<sub>3</sub>O and 40% CdS<sub>4</sub> <sup>d</sup> from ref 23 <sup>e</sup> this peptide is 100% CdS<sub>3</sub>O <sup>f</sup> from ref 25 <sup>g</sup> from ref. 9