Protein-Cadmium Interactions in Crowded Biomolecular Environments Probed by In-cell and Lysate NMR Spectroscopy

Sachin S. Katti[a] and Tatyana I. Igumenova*

Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas 77843, United States

* Tatyana.Igumenova@ag.tamu.edu

[a] Current address: Department of Cell Biology & Genetics, Texas A&M Health Science Center, Texas A&M University, College Station, Texas, United States

Abstract

One of the mechanisms by which toxic metal ions interfere with cellular functions is ionic mimicry, where they bind to protein sites in lieu of native metals Ca^{2+} and Zn^{2+}. The influence of crowded intracellular environments on these interactions is not well understood. Here, we demonstrate the application of *in-cell* and lysate NMR spectroscopy to obtain atomic-level information on how a potent environmental toxin cadmium interacts with its protein targets. The experiments, conducted in intact *E. coli* cells and their lysates, revealed that Cd^{2+} can profoundly affect the quinary interactions of its protein partners, and can replace Zn^{2+} in both labile and non-labile protein structural sites without significant perturbation of the membrane binding function. Surprisingly, in crowded molecular environments Cd^{2+} can effectively target not only all-sulfur and mixed sulfur/nitrogen but also all-oxygen coordination sites. The sulfur-rich coordination environments show significant promise for bioremedial applications, as demonstrated by the ability of the designed protein scaffold α_3DIV to sequester intracellular cadmium. Our data suggests that *in-cell* NMR spectroscopy is a powerful tool for probing interactions of toxic metal ions with their potential protein targets, and for the assessment of potency of sequestering agents.

Keywords: Cadmium • In-cell NMR • Protein design • Protein kinase C • Quinary Interactions

Introduction

Nutritive, or native, metal ions are essential for life-sustaining biological processes.[1] Metal ion-binding events affect the structural and functional properties of biological macromolecules, thereby regulating catalysis,[2] signal transduction,[3] and energy generation.[4] The selectivity of proteinic metal-ion interaction sites is defined by the structure of the first coordination shell and Lewis acidity of the ligands.[5] Xenobiotic metals can subvert these selectivity barriers through ionic mimicry, where they directly compete with native metal ions for the binding sites in biological macromolecules. These non-native interactions interfere with vital cellular processes whose dysregulation poses significant threat to human health. Among notable examples, chronic exposure to potent environmental toxin Cd^{2+} is associated with neurological, respiratory, renal, and reproductive disorders.[6]

The structural analysis of Cd^{2+}-bound proteins in the Protein Data Bank (PDB) reports that Cd^{2+} is found in a variety of coordination environments comprising O, S, and N ligands, and can adopt coordination numbers ranging from 4 to 8.[7] Consistent with this finding, several *in vitro* studies have demonstrated the ability of Cd^{2+} to effectively displace Zn^{2+} and Ca^{2+} from the thiol- and oxygen-rich sites of proteins and protein domains.[8] However, Cd^{2+} interactions in cellular environments can be affected by several factors such as influx-efflux homeostasis, accessibility of target sites, nutritive metal-ion pools, and intracellular chelators.[9] To establish biological relevance, specificity and affinity of Cd^{2+} towards putative protein targets need to be examined in environments reflecting these conditions.

In crowded environments, detection of protein spectra and binding events requires advanced isotope enrichment schemes and rapid multidimensional data acquisition methods.[10] Notable demonstrations of *in-cell* NMR spectroscopy applied to the studies of protein-metal ion

42 interactions include Zn^{2+} binding to Cu/Zn superoxide dismutase 1 in HEK293T cells,[11] and
43 characterization of Ca^{2+}-induced conformational rearrangement of calmodulin in *Xenopus laevis*
44 oocytes.[12] We have previously demonstrated that in buffered solutions, amide ^{1}H and ^{15}N NMR
45 chemical shifts in proteins are exquisitely sensitive to the changes in electrostatic environment
46 imparted by Cd^{2+} binding, leading to the unambiguous identification of Cd^{2+} binding sites.[8b-e]
47 The design of *in-cell* NMR studies to probe protein-Cd^{2+} interactions requires a live cell
48 system resistant to heavy metals. *E. coli* cells exhibit substantial tolerance towards heavy metals
49 like Cd^{2+}, which is attributed to their evolved and adaptable resistance mechanisms.[13] In this
50 work, we present an approach to investigate protein-Cd^{2+} interactions in crowded biomolecular
51 environments provided by *E. coli* cells and lysates (**Figure 1A**). In this approach, Cd^{2+} (or any
52 other divalent metal ion) is added directly to the *E. coli* cell culture transformed with a plasmid
53 that carries the target protein gene under the control of an inducible promoter. Addition of metal
54 ions takes place after cells are transferred to the isotopically enriched growth medium and is
55 timed to the induction of protein expression. Using 2D ^{1}H-^{15}N chemical shift correlation NMR
56 spectroscopy, the effect of metal ions on the target protein is monitored directly in intact cells
57 and lysates. Here, we apply this approach to three metal-binding protein systems to evaluate their
58 susceptibility to Cd^{2+} interactions, establish the effect of Cd^{2+} on the quinary structure, and
59 assess their potential as sequestering agents and sensors of cellular Cd^{2+} uptake.
60
61
62
63
64

Results and discussion

Choice of Protein Systems

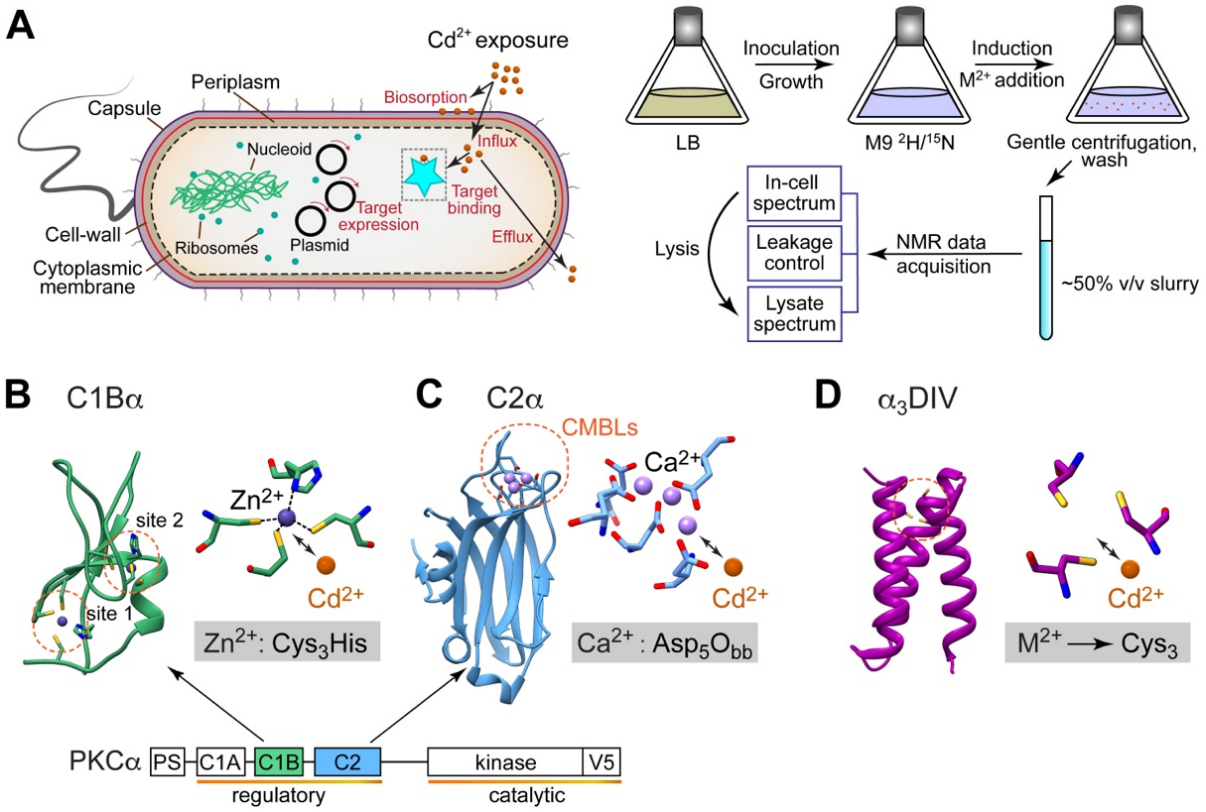


Figure 1. Overview of the experimental approach and protein systems for Cd^{2+} studies. *(A) Experimental design for the detection of protein-Cd^{2+} interactions in intact E. coli cells and their lysates. NMR samples contain a suspension of intact E. coli cells that express target proteins and are grown in the isotopically-enriched medium with M^{2+} (M=Cd, Zn, Ca) supplementation. (B-D) Three-dimensional structures and metal-ion coordination sites of (B) C1Bα (PDB:2ELI), (C) C2α (PDB:3GPE), and (D) α3DIV (PDB:2MTQ). The modular diagram of PKCα shows regulatory and catalytic regions, with C1Bα and C2α domains color-coded green and blue, respectively.*

The three protein systems chosen for this study present different metal-ion coordination environments to probe Cd^{2+} interactions, and have different folds, ligand preferences, and electrostatic surface properties. The first two proteins are independently folded peripheral membrane-binding modules of the conventional α isoform of protein kinase C (PKCα): the second cysteine-rich conserved homology-1 domain (C1Bα) and the conserved homology-2 domain (C2α) (**Figure 1B-C**). PKCα is a Ca^{2+}-dependent Ser/Thr protein kinase of the AGC

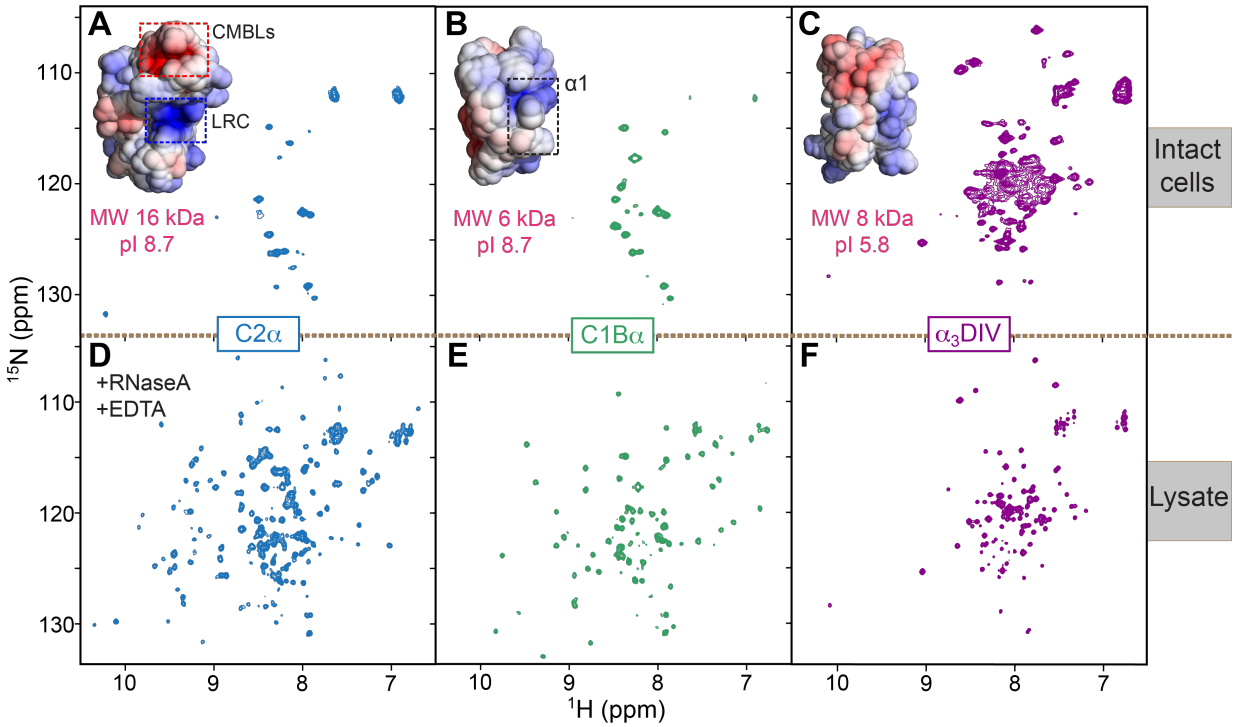
82 family. This enzyme is a regulatory node of the GPCR-mediated phosphoinositide signaling
83 pathway and a proposed molecular target of Cd^{2+}.[14] Activation of PKCα requires membrane
84 recruitment of its regulatory region comprising C1 and C2 domains, both of which contain metal
85 ion binding sites that could be targeted by Cd^{2+} in the cellular environment. In C1Bα, these are
86 represented by two Cys_3His motifs, each coordinating a structural Zn^{2+} ion (**Figure 1B**).[15] In
87 C2α, it is the aspartate-rich Ca^{2+}- and membrane-binding loops (CMBLs) that are of interest
88 (**Figure 1C**).[16] The third protein system is an engineered heavy metal-binding peptide, α_3DIV,
89 designed by Pecoraro's laboratory (**Figure 1D**).[17] This synthetic helical scaffold contains a pH-
90 dependent, self-organizing Cys_3 motif that mimics thiol-rich sites of metalloproteins, such as
91 metallothionines.[18] We and others have demonstrated all three proteins bind Cd^{2+} with high-
92 affinity in aqueous buffer solutions.[8b, 17]
93
94 **NMR Detection of Target Proteins in Intact Cells**
95 The first step was to establish the feasibility of detecting Cd^{2+}-free C1Bα, C2α, and
96 α_3DIV in intact cells and lysates using 2D [^{1}H,^{15}N] NMR spectroscopy. The general sample
97 preparation workflow involved recombinant expression of target proteins in *E. coli*, with cell
98 cultures grown on the M9 minimal medium supplemented with [^{15}N-enriched] NH_4Cl. The
99 growth medium was prepared in 100% D_2O to achieve protein deuteration and thereby reduce
100 NMR peak linewidths by attenuating ^{1}H- mediated cross-relaxation pathways.
101 The 2D [^{15}N-^{1}H] SOFAST-HMQC NMR spectra collected on the 50% v/v slurry of intact
102 cells show significant broadening of the ^{15}N-^1H_N resonances (**Figures 2A-C**). This is consistent
103 with the restricted molecular tumbling in a high-viscosity cellular milieu, caused by proteins
104 engaging in transient, low-affinity "quinary" interactions with endogenous macromolecules.[19]

105 Control NMR spectra on supernatants of cell suspensions were recorded after each *in-cell*

106 experiment to ensure cell integrity and no protein leakage (representative spectrum is shown in

107 **Figure S1A**).

108



109

110 ***Figure 2. 2D NMR detection of target proteins in intact E. coli cells and lysates.*** *The in-cell/lysate [^{15}N-*
111 *^{1}H] SOFAST-HMQC spectral pairs are shown for C2α (A,D), C1Bα (B,E), and α$_3$DIV (C,F). The protein*
112 *surfaces are color-coded according to the APBS-calculated electrostatic potential (± 5kT/e, insets of A-*
113 *C). CMBL and LRC regions of C2α, and the α1 helix of C1Bα are highlighted with dashed boxes in (A)*
114 *and (B), respectively. The amide chemical shift pattern in the in-cell α$_3$DIV spectrum (C) is characteristic*
115 *of an α-helical fold. Significant broadening of the amide resonances is observed for C2α and C1Bα in*
116 *intact cells. All lysate spectra (D-F) report on fully folded proteins.*
117

118 We find that the *in-cell* α$_3$DIV spectrum contains many amide cross-peaks whose overall

119 chemical shift dispersion and pattern report on a well-folded helical structure (**Figure 2C**). In

120 contrast, the C2α and C1Bα data contain only a small subset of protein and metabolite cross-

121 peaks in a narrow spectral range centered around ~8.3 ppm (**Figure 2A, B**). To evaluate if C2α

122 and C1Bα are properly folded, we conducted NMR experiments on the clarified lysates

123 generated from the *in-cell* NMR samples. Persistent molecular crowding in lysates was verified
124 by SDS-PAGE (**Figure S1B**).
125
126 **NMR Detection of Target Proteins in Lysates**
127 Significant line-broadening persisted in the C2α spectrum even after the cell lysis
128 (**Figure S1C**). However, addition of both EDTA and RNase A resulted in the appearance of
129 well-dispersed ^{15}N-$^{1}H_N$ cross-peaks indicative of a fully folded protein (**Figure 2D**). The line-
130 broadening can therefore be attributed to the C2α interactions with metal ions and RNA[20] in the
131 crowded environment. The Ca^{2+}-binding site of C2α resides on the apical CMBL region (**Figure**
132 **1C**) that can potentially interact with abundant divalent metal ions such as Mg^{2+}. Indeed, only
133 RNase A treatment was required to recover full spectra when C2α-expressing cells were grown
134 in the presence of 50 μM EDTA and reduced Mg^{2+} levels. We hypothesized that the C2α region
135 responsible for RNA interactions is the electropositive lysine-rich cluster, or LRC (**Figures 2A,**
136 **S2A**) that mediates PKC recruitment to membranes through interactions with negatively charged
137 lipids such as phosphatidylserine (PtdSer) and phosphatidylinositol-4,5-bisphosphate
138 (PtdIns4,5P_2).[16a, 21] When we mutated two of the four Lys residues of the LRC to Glu, EDTA-
139 only treatment was sufficient to obtain a well-resolved spectrum (**Figure S2A-B**) indicating that
140 interactions with RNA are attenuated in this charge-reversal KEKE C2α mutant. In contrast to
141 C2α, the lysate of cells expressing C1Bα did not require addition of RNase A or EDTA to
142 produce a well-dispersed chemical shift pattern that reports on the fully folded protein (**Figure**
143 **2E**).
144 Pairwise comparison of the *in-cell* and lysate spectra for C1Bα and C2α (**Figures 2A-D,**
145 **2B-E, S1D**) suggests that line-broadening in the *in-cell* data is caused by both proteins engaging

146 in significant quinary interactions with intracellular components. The propensity to engage in
147 quinary interactions correlates with the existence of distinct electropositive regions, as both
148 proteins have an isoelectric point of 8.7. In C1Bα, the relevant region is likely to be the α1 helix
149 that is responsible for PtdSer recognition at the membranes (**Figure 2E**).[15a] In C2α it is the
150 CMBLs and the LRC regions whose interactions with intracellular binding partners might be
151 enhanced in much the same way as Ca^{2+} binding facilitates the PKC interactions with negatively
152 charged phospholipids. We note that in contrast to C1Bα and C2α, α_3DIV is acidic (pI=5.8) and
153 its electrostatic surface presents a rather disperse distribution of positively and negatively
154 charged regions. As a result, α_3DIV does not engage in any significant quinary interactions,
155 evident by the comparable quality of the *in-cell* and lysate spectra (**Figure 2C-F**). Collectively,
156 our data recommend the use of lysates to probe Cd^{2+} interactions with PKCα regulatory
157 domains, and intact cells to probe Cd^{2+} interactions with α_3DIV.
158
159 **Cd^{2+} Substitutes for Zn^{2+} in C1Bα Without Loss of Ligand-sensing Function**
160 Structural Zn^{2+} ions, each coordinated in a tetrahedral geometry by a Cys_3His motif are essential
161 for the stability of the C1Bα fold (**Figure 3A**).[15a] The affinity of Zn^{2+} to thiol-rich structural protein sites
162 is estimated to be in the pico-to-femtomolar range. To determine if C1Bα incorporates Cd^{2+} in the
163 presence of native Zn^{2+}, we expressed the protein with 25 μM Zn^{2+} and 50 μM Cd^{2+} added to the *E. coli*
164 culture. The [^{15}N-^{1}H] SOFAST-HMQC spectrum of cell lysate differs from that obtained in the Zn^{2+}-only
165 control experiment (**Figure 3A**). Notably, residues located at and near the Zn^{2+} coordination sites 1 and 2
166 show additional subset of amide cross-peaks whose position matches those previously obtained for the
167 Cd^{2+}-treated protein in an aqueous buffer.[8e] We conclude that in the crowded environment Cd^{2+}
168 successfully competes with Zn^{2+} for the Cys_3His motifs of C1Bα and supports its three-dimensional fold.
169

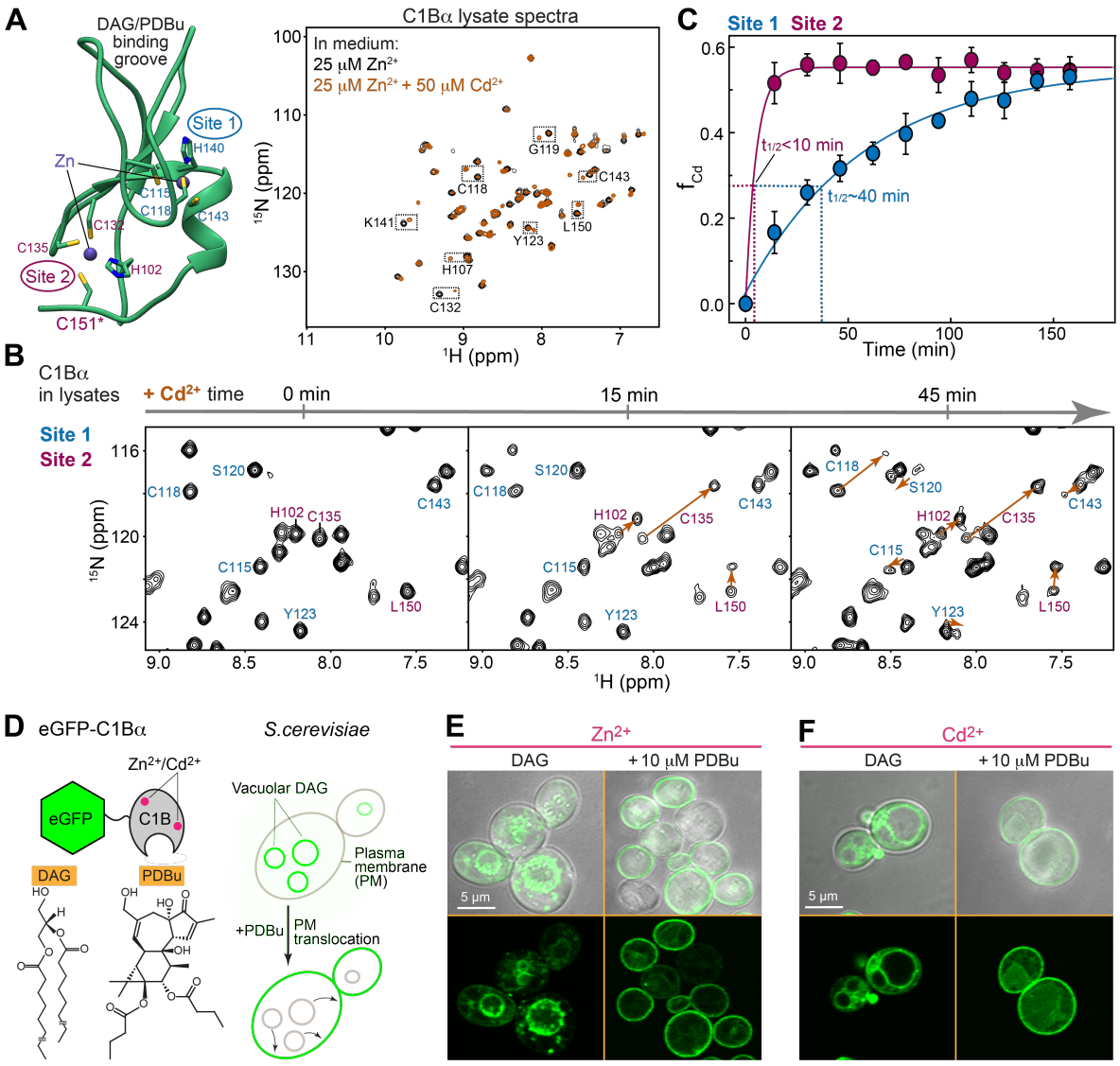


Figure 3. Displacement of structural Zn^{2+} by Cd^{2+} generates a functionally competent C1Bα fold. *(A) Cys_3His structural Zn^{2+} sites of C1Bα (left panel, PDB:2ELI) and the spectroscopic signatures of their $Zn^{2+}\rightarrow Cd^{2+}$ substitution (right panel). The latter is illustrated by the superposition of [^{15}N-^{1}H] SOFAST-HMQC lysate spectra of cells expressing C1Bα in the presence of Zn^{2+} (black) and Zn^{2+}-Cd^{2+} mixture (brown). Cd^{2+} binding to C1Bα results in the appearance of distinct amide resonances for several residues (highlighted by squares). (B,C) Kinetics of $Zn^{2+}\rightarrow Cd^{2+}$ substitution initiated by the addition of 2-fold molar excess of Cd^{2+} to Zn^{2+}-bound C1Bα. (B) shows expansions of representative [^{15}N-^{1}H] SOFAST-HMQC lysate spectra collected at 3 time points. The cross-peak labels of residues responsive to Cd^{2+} binding at sites 1 and 2 are color-coded blue and maroon, respectively. The arrows connect cross-peaks that belong to the same residue in Zn^{2+}- and Cd^{2+}-bound C1Bα species. (C) Time course of $Zn^{2+}\rightarrow Cd^{2+}$ substitution reported by the mean fractional population of Cd^{2+} bound C1Bα (f_{Cd}). Solid lines are to guide the eye. $t_{1/2}$ are estimated as the time needed to reach 50% of the plateau f_{Cd} values. (D) Chemical structures of C1Bα ligands and experimental design of fluorescence microscopy experiments in S. cerevisiae. C1Bα is fused with eGFP to enable detection of intracellular DAG-dependent localization and translocation to plasma membrane upon stimulation with phorbol 12,13-dibutyrate (PDBu). (E,F) Live-cell confocal microscopy images of S. cerevisiae cells grown in SD medium (-URA) supplemented*

187 *with either 10 μM Zn^{2+} (E) or Cd^{2+} (F). The membrane localization pattern of eGFP-C1Bα is similar*
188 *when Cd^{2+} is added to the medium instead of Zn^{2+}.*
189

190 To establish the potency of Cd^{2+} in displacing Zn^{2+} from C1Bα, we monitored the metal ion
191 substitution reaction by NMR spectroscopy in real time. The experiment involved adding purified 0.2
192 mM Zn^{2+}-containing [U-^{15}N] enriched C1Bα to the concentrated lysate of *E. coli* with no expression
193 plasmid; initiating the $Zn^{2+}\rightarrow Cd^{2+}$ substitution reaction by adding two-fold molar excess of Cd^{2+}; and
194 collecting [^{15}N-^{1}H] SOFAST-HMQC spectrum every ~15 minutes. We find that the formation of C1Bα
195 species with Cd^{2+} at site 2, C1B$\alpha_{Cd(2)}$, occurs within the first 15 minutes and is manifested by the
196 appearance of characteristic amide cross-peaks of His102, Cys135, and Leu150 (**Figure 3B**). A
197 considerably slower formation of C1Bα species with Cd^{2+} at site 1, C1B$\alpha_{Cd(1)}$, is reported by the
198 appearance of characteristic amide cross-peaks of Cys118, Cys143, and Cys115.
199 To evaluate the relative susceptibility of Zn^{2+}-coordinating Cys_3His motifs to Cd^{2+} substitution,
200 we calculated the fractional populations f_{Cd} of C1B$\alpha_{Cd(1)}$ and C1B$\alpha_{Cd(2)}$ species using NMR peak
201 intensities and analyzed their accumulation kinetics. Site 2 is considerably more susceptible to Cd^{2+} ($t_{1/2}$ <
202 10 min) than site 1 ($t_{1/2}$ ~ 40 min) (**Figure 3C**). This agrees well with our previous findings that site 2 of
203 C1Bα is dynamic and prone to oxidative modifications.[22] The dynamics is associated with the loss of the
204 Zn-S(Cys151) coordination bond (**Figure 3A**) and the concomitant local conformational change of the
205 protein. The latter increases solvent exposure of the Cys132-Cys135-Cys151-His102 motif, making it
206 more susceptible to Cd^{2+} invasion. Of note, both kinetic curves shown in **Figure 3C** plateau at the f_{Cd}
207 values of ~55%, suggesting that in lysates the affinities of sites 1 and 2 to Cd^{2+} are comparable.
208 How does Cd^{2+} substitution affect the lipid-sensing function of C1Bα? We chose two ligands to
209 probe the C1Bα function in live cells: (i) *sn*-1,2-diacylglycerol (DAG), a native lipid activator of PKC;
210 and (ii) phorbol 12,13-dibutyrate (PDBu), a potent tumor-promoting activator of PKC (**Figure 3D**). Our
211 recent structural work on the C1::ligand complexes[15b] demonstrates how the integrity of the C1B domain
212 and particularly of its apical loop region (**Figure 3A**) are essential for the recognition and capture of its

213 membrane-embedded ligands. To evaluate the effect of Cd^{2+} on the C1Bα function, we monitored the
214 ligand-dependent localization of the eGFP-C1Bα fusion protein in live *S. cerevisiae* cells using
215 fluorescence microscopy (**Figure 3D**). *S. cerevisiae* cells were chosen because they provide a eukaryotic
216 environment with high basal level of endogenous DAG.

217 We find that in the presence of 10 μM Zn^{2+} eGFP-C1Bα preferentially localizes to vacuolar
218 rather than plasma membrane DAG pools (the plasma membrane is marked by the cell boundaries in
219 bright field images shown in **Figure 3E-F**). The membrane localization pattern is similar when Cd^{2+} is
220 added to the medium instead of Zn^{2+}, indicating that Cd^{2+}-substituted C1Bα retains DAG sensitivity.
221 Treatment of cells with PDBu, a high-affinity competitor of DAG, results in essentially quantitative
222 translocation of eGFP-C1Bα to the plasma membrane. That response was identical for cells exposed to
223 Zn^{2+} or Cd^{2+}. The appearance of cytoplasmic puncta were observed only when cells were treated with
224 higher concentrations of Cd^{2+} (25 μM in medium, **Figure S3**). These puncta were frequently associated
225 with the limiting membrane of the vacuole, and likely correspond to the fraction of eGFP-C1Bα
226 subspecies that misfold and aggregate. The properly folded subspecies remain ligand-responsive and
227 translocate to the plasma membrane upon PDBu treatment (**Figure S3**).

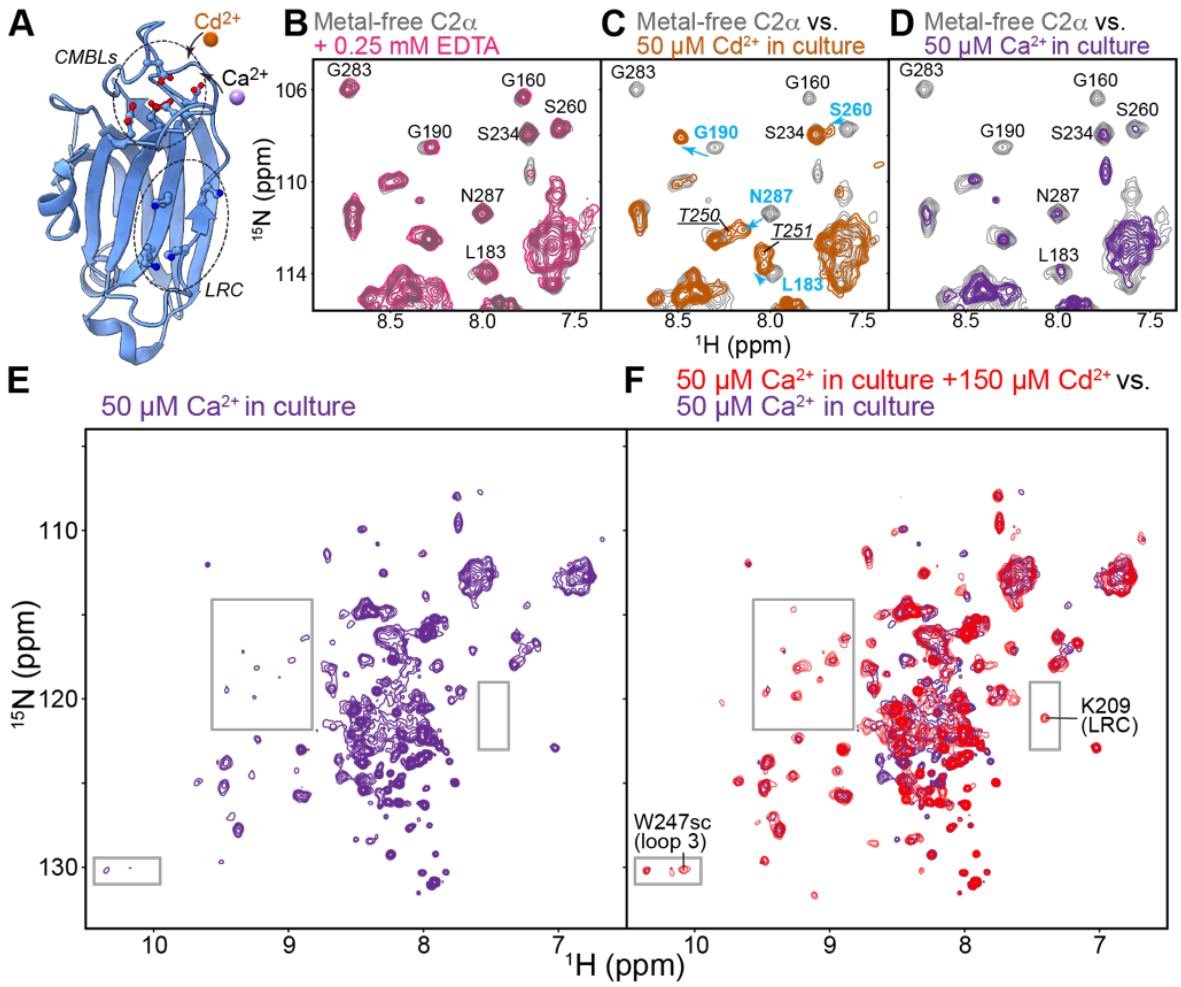
228 In summary, we established that Cd^{2+} targets high-affinity structural Zn^{2+} sites of C1Bα leading
229 to the formation of Cd^{2+}-substituted C1Bα that retains its ligand-sensing function. To our knowledge, the
230 only other reported case of "isofunctional" Cd^{2+} substitution of Zn^{2+} in proteins is the Cadmium carbonic
231 anhydrase of marine diatoms, which retains its catalytic activity as a Cd^{2+}-bound complex.[23] Our
232 findings indicate that C1 domains of other PKC isoforms as well as those of other diacylglycerol effector
233 proteins are potential targets of Cd^{2+}. The co-occurrence of Zn and Cd-folds in C1Bα is analogous to how
234 metallothionines incorporate Cd^{2+} in mixed Zn-Cd thiolate clusters.[24] Given that many transcription
235 factors bind Zn^{2+} via zinc finger motifs, and that some of these DNA-binding proteins execute Zn^{2+}
236 sensing functions, there is building evidence to indicate that Cd^{2+} substitution in these cases contributes to

237 Cd^{2+} toxicity.[25] In this regard, as Zn^{2+} binding is a key feature of a large number of proteins in addition to

238 transcription factors, the potential target space for Cd^{2+} substitution is large.[26]

239

240 **C2α Interacts with Cd^{2+} in *E. coli* Lysates**



241

242 ***Figure 4. C2α binds Cd^{2+} and Ca^{2+} in cell lysates.*** *(A) Three-dimensional structure of C2α (PDB:3GPE)*
243 *showing the location of Ca^{2+}- and membrane-binding loops (CMBLs) and the Lysine-Rich Cluster (LRC).*
244 *CMBLs bind Cd^{2+}, while LRC interacts with anionic species. (B-D) Expansions of [^{15}N-^{1}H] SOFAST-*
245 *HMQC spectra collected on lysates of E. coli cells expressing C2α in the presence of 10 μM Mg^{2+} (B), 50*
246 *μM Cd^{2+} (C), and 50 μM Ca^{2+} (D) in the medium. All lysates were treated with RNase A prior to data*
247 *collection. In (B), the C2α spectrum in Mg^{2+}-containing lysate (gray) does not appreciably change upon*
248 *EDTA treatment (pink), indicating that it corresponds to the metal-free state of the protein. In (C), several*
249 *C2α residues in Cd^{2+}-containing lysate (brown) show chemical shift changes characteristic of the Cd^{2+}-*
250 *bound protein. The changes are shown with blue arrows. The cross-peaks of the underlined residues T250*
251 *and T251 are broadened in the apo C2α but reappear upon Cd^{2+} binding. In (D), the C2α spectrum in*
252 *Ca^{2+}-containing lysate (purple) shows significant linebroadening compared to the metal-free protein*
253 *(gray). (E) Full [^{15}N-^{1}H] SOFAST-HMQC spectrum of C2α in Ca^{2+}-containing lysate (purple), and (F)*

254 *its overlay with spectrum obtained upon addition of 150 μM Cd^{2+} into the lysate (red). Cd^{2+} addition*
255 *results in the recovery of resonance intensities, evident in the spectral regions highlighted by rectangles.*
256 *W247 and K209 are noted as representative residues for loop 3 and LRC regions, respectively.*
257

258 C2α binds its native metal ion, Ca^{2+}, via an all-oxygen coordination environment
259 provided by 5 conserved aspartate residues of the CMBL region (**Figure 4A**). It can
260 accommodate up to 3 Ca^{2+} ions whose affinities fall into the μM-mM range. Our previous work
261 has established, using biophysical measurements and structural data, that in aqueous buffer
262 solutions Cd^{2+} binds to C2α with comparable stoichiometry and >30-fold higher affinity.[8b]
263 Since oxygen is a harder Lewis base than sulfur, Cd^{2+} affinity to all-oxygen coordination sites is
264 generally weaker than that to all-sulfur sites. To determine if C2α-Cd^{2+} interactions take place in
265 the crowded molecular environment that contains many potential Cd^{2+} chelators, we conducted
266 experiments on lysates of cells that express isotopically enriched C2α. To generate a clean
267 metal-free reference state of C2α, we expressed the protein in the presence of 50 μM EDTA in
268 the medium under limiting (10 μM) Mg^{2+} supplementation. The [^{15}N-^{1}H] SOFAST-HMQC
269 spectrum of RNase A-treated lysate remained unchanged upon EDTA addition and showed well-
270 dispersed cross-peaks characteristic of a fully folded protein (**Figure 4B**).
271 Next, we prepared lysate samples of cells grown with either 50 μM Cd^{2+} or Ca^{2+} but no
272 EDTA added in culture. The chemical shifts of the C2α amides in Cd^{2+}-containing lysate clearly
273 report on the formation of the C2α-Cd^{2+} complex[8b] (**Figure 4C**). For instance, the chemical
274 shift perturbations of Gly190, Asn287, and Leu183 amide resonances are similar to those
275 detected in C2α upon Cd^{2+} binding in aqueous buffer.[8b] The cross-peaks of Thr150 and Thr151
276 of CMBL3 that are exchange-broadened beyond detection in the metal-free state, are readily

277 detectable in the Cd^{2+}-containing lysate due to the rigidification of the CMBL region that takes
278 place upon metal ion complexation (**Figure 4C**).

279 In sharp contrast with Cd^{2+}, the C2α amide peaks in the Ca^{2+}-containing lysate are
280 significantly broadened (**Figures 4D-E**). Since this broadening affected most protein residues
281 and not just those of the Ca^{2+}-binding region, we conclude that Ca^{2+}-complexed C2α engages
282 extensively in quinary interactions. Indeed, Ca^{2+}-induced NMR line broadening is reversible by
283 EDTA treatment. When Cd^{2+} is added to the Ca^{2+}-containing lysate many of the protein cross-
284 peaks recover their intensities (**Figure 4F**). The latter finding suggests that Cd^{2+} can displace
285 native Ca^{2+} from C2α and thereby affect its propensity to engage in quinary interactions. One
286 plausible explanation for this are coordination preferences of Ca^{2+} and Cd^{2+}. In biological
287 macromolecules, Ca^{2+} has a potential to expand its coordination number up to 8, whereas for
288 Cd^{2+} the coordination numbers of 4 to 6 are more common.[27] Therefore, Ca^{2+} might directly
289 mediate the interactions of C2α with polyanionic compounds within cells/lysates. This
290 interpretation is supported by our previous findings that Ca^{2+}-bound C2α and both C2 domains
291 of Synaptotagmin 1 readily associate with anionic membranes (specifically anionic lipids PtdSer
292 and PtdIns(4,5)P_2) as part of their signal transduction mechanism, while Cd^{2+}-bound proteins are
293 unable to bind PtdSer-containing membranes.[8b, 8c]

294 In summary, we have demonstrated that Cd^{2+} is capable of targeting oxygen-rich sites
295 and displacing native metal ion Ca^{2+} from C2α. In a broader context, aspartate and glutamate
296 side chains frequently serve as bidentate ligands in the selectivity filters of ion channels and
297 metal-ion sensing EF-hand proteins. The NMR-based approach presented here can be applied to
298 test the Cd^{2+} susceptibility of such protein targets in the crowded environment.
299

α_3DIV Sequesters Cd^{2+} in *E. coli* Cells

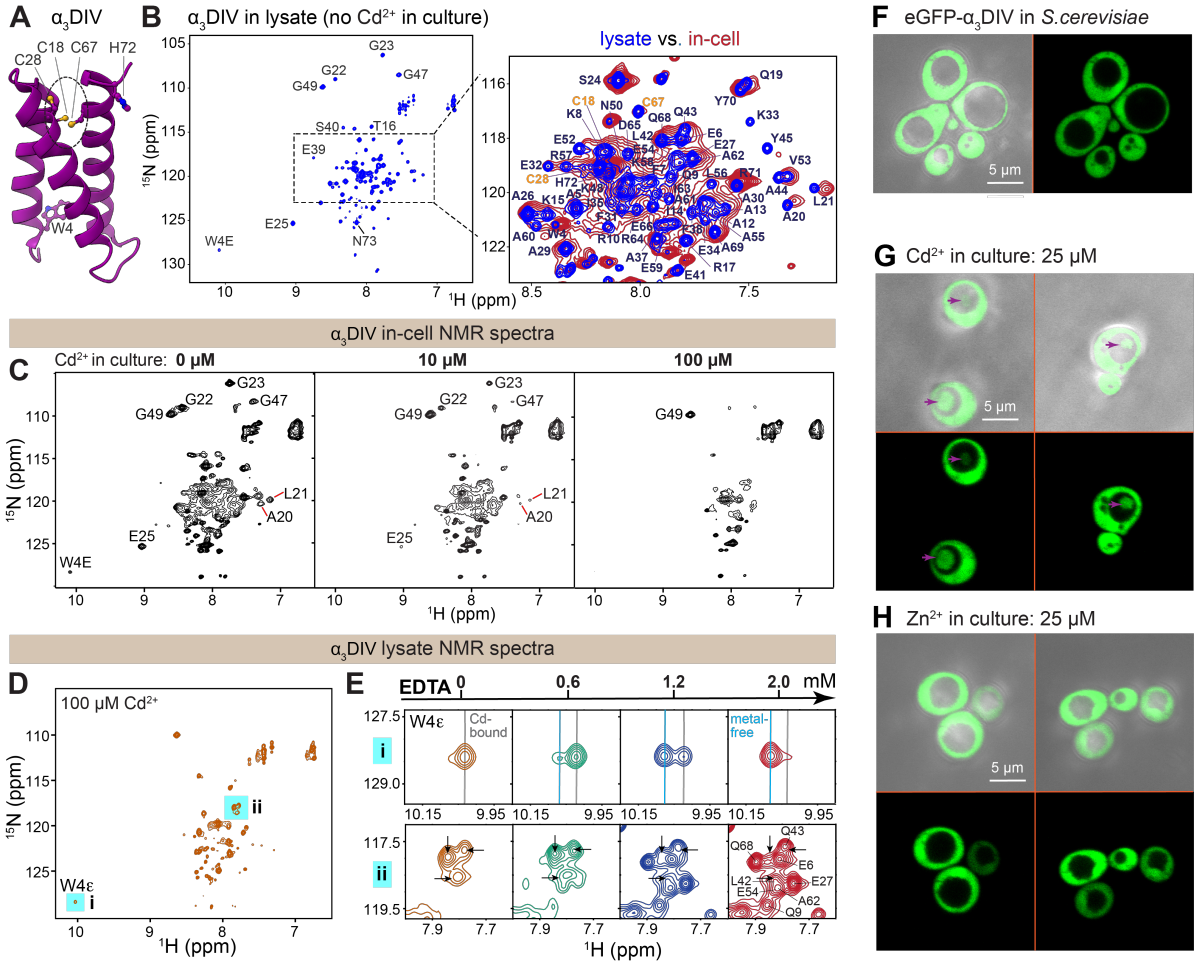


Figure 5. α_3DIV sequesters Cd^{2+} in cells. *(A) Three-dimensional structure of α_3DIV (PDB:2MTQ) highlighting Cd^{2+} coordination motif formed by three Cys residues: 18, 28, and 67. His72 may also be involved in metal ion coordination. (B) [^{15}N-^{1}H] SOFAST-HMQC spectrum of α_3DIV in E. coli Cd^{2+}-free lysate (blue, left panel) and its overlay with the in-cell spectrum (red, right panel). (C) In-cell [^{15}N-^{1}H] SOFAST-HMQC spectra of α_3DIV at 0, 10, and 100 μM Cd^{2+} added in culture. Progressive line-broadening is observed upon increasing Cd^{2+} concentration. (D) [^{15}N-^{1}H] SOFAST-HMQC spectrum of α_3DIV in lysate generated from cells grown with 100 μM Cd^{2+} in culture. Addition of increasing EDTA concentration to this sample results in the recovery of apo α_3DIV resonances, as shown by the adjacent spectral expansions of regions i and ii in (E). (F) Live-cell confocal microscopy image of S. cerevisiae cells expressing eGFP-α_3DIV grown in SD medium (-URA). The protein shows a diffuse cytosolic localization pattern. (G) Addition of Cd^{2+} to the growth medium results in the import of detectable eGFP-α_3DIV fraction (marked with pink arrows) into the vacuolar lumen. (H) Addition of Zn^{2+} has no significant effect on the cytosolic distribution of eGFP-α_3DIV. In (F-H), bright field images mark the cell boundary.*

317 α_3DIV is a designed metal-ion sensing protein of 73 amino acids that contains a tri-
318 cysteine motif (**Figure 5A**). This motif binds heavy metal ions Cd^{2+}, Pb^{2+}, and Hg^{2+} with sub-
319 μM affinities in aqueous solutions.[17] The metal-ion binding site of α_3DIV resembles thiol-rich
320 coordination sites found in bacterial MerR family of transcriptional regulators[8a, 28] that play key
321 roles in heavy-metal efflux and detoxification processes. While no high-resolution metal-ion
322 bound structure of α_3DIV is available, ^{113}Cd NMR and ^{111m}Cd PAC spectroscopy measurements
323 suggest that Cd^{2+} adopts a pseudo-tetrahedral geometry $CdS_3(O/N)^-$, where O belongs to the
324 exogenous water molecule and N belongs to His72 of α_3DIV.[17] We sought to evaluate the
325 potential of α_3DIV as a Cd^{2+}-sensing and sequestering protein in the cellular environment.
326 The ability of α_3DIV to bind heavy metal ions relies on the pre-organization of the tri-
327 cysteine motif through the formation of the triple-helix bundle. This assembly process is pH
328 dependent.[18] To evaluate the α_3DIV state in the crowded environment, we mapped the amide
329 cross-peak assignments obtained in aqueous buffer[18] onto the [^{15}N-^{1}H] SOFAST-HMQC
330 spectra of Cd^{2+}-free lysates and intact cells expressing α_3DIV (**Figures 5B, C**). Based on the
331 excellent agreement between the α_3DIV chemical shifts in all three environments (buffer, lysate,
332 and intact cells), we conclude that α_3DIV maintains its folded state with the pre-organized tri-
333 cysteine motif in the cellular milieu.
334 To determine if α_3DIV can serve as a cellular reporter of Cd^{2+} uptake, we expressed the
335 protein in *E. coli* with 10 and 100 μM Cd^{2+} added in culture. Comparison of the *in-cell* NMR
336 spectra of α_3DIV at 0, 10 and 100 μM Cd^{2+} shows significant peak broadening as Cd^{2+}
337 concentration increases (**Figure 5C**). Although we cannot exclude the engagement of α_3DIV in
338 quinary interactions, the most likely cause is the chemical exchange of Cd^{2+}-bound α_3DIV that is

339 intermediate on the chemical shift timescale. In fact, a similar pattern of line-broadening was
340 observed previously in the NMR spectra of Pb^{2+}- and Hg^{2+}-bound α_3DIV (aqueous buffer), and
341 was attributed to structural dynamics associated with metal ions adopting distinct coordination
342 geometries within the asymmetric tri-cysteine site.[18] Not surprisingly, Cd^{2+}-induced peak
343 broadening persists in cell lysates (**Figures 5D and S4**). Treatment of lysates with increasing
344 concentrations of EDTA removes bound Cd^{2+} from α_3DIV and results in the recovery of
345 resonances that correspond to the metal-free α_3DIV (**Figures 5E, S4A-B**). This is aptly
346 illustrated by the behavior of the ^{15}Ne-^{1}He Trp4 cross-peak that has distinct ^{1}H chemical shifts
347 for the apo and Cd^{2+}-bound α_3DIV species (**Figures 5E, S4C**). Together, these findings
348 demonstrate that α_3DIV sequesters Cd^{2+} taken up by *E. coli*.

349 Next, we tested the specificity of α_3DIV as a Cd^{2+} sensor in both prokaryotic and
350 eukaryotic environments by introducing Zn^{2+} as a competing metal ion. The NMR spectrum of
351 cells grown with 100 μM Zn^{2+} added in culture revealed only insignificant changes compared to
352 the no-metal ion control, suggesting that no specific α_3DIV-Zn^{2+} interactions take place in the
353 cellular environment (**Figure S5A**). Very high concentrations of Zn^{2+}, 0.4 mM, added externally
354 to cell lysates were required to detect minor changes in the NMR spectra (**Figure S5B**), further
355 supporting our conclusion that no specific high-affinity interactions with Zn^{2+} take place.

356 We assessed the Cd^{2+}- and Zn^{2+}-dependent behavior of α_3DIV in eukaryotic environment
357 by monitoring the localization of the eGFP-α_3DIV in *S.cerevisiae* using fluorescence
358 microscopy. In the absence of direct metal ion supplementation, eGFP-α_3DIV shows a cytosolic
359 distribution (**Figure 5F**). Upon addition of Cd^{2+} to the concentration of 25 μM, however, a
360 detectable fraction of eGFP-α_3DIV is imported into to the vacuolar lumen (**Figure 5G**). This
361 relocalization is not observed upon addition of 25 μM Zn^{2+} (**Figure 5H**). We interpret these data

362 to indicate that Cd^{2+}-bound α_3DIV is a substrate for direct transport from the cytoplasm into the
363 vacuole via the cytoplasm-to-vacuole targeting (Cvt) pathway.[29] We speculate eGFP-α_3DIV
364 has a propensity to aggregate and that it is this aggregation that induces its clearance from the
365 cytoplasm.

366 In summary, we have demonstrated that α_3DIV responds to cellular Cd^{2+} levels by
367 showing characteristic NMR signatures and localization patterns. A combination of robust
368 structural framework provided by the α_3D triple-helix bundle and the engineered metal-ion
369 binding sites provides an opportunity to generate a variety of metal sensors and chelators.[30]
370 Potential applications of these protein scaffolds include bioremediation using microbial systems,
371 investigation of heavy metal resistance mechanisms, and rational design of coordination
372 environments to achieve high metal selectivity.[31] *In-cell* NMR and direct imaging methods, as
373 demonstrated for the α_3DIV-Cd^{2+} system, show significant promise in the evaluation of metal-
374 ion binding properties in the crowded environments and in guiding further design efforts.
375

376 **Conclusions**

377 Metal binding properties of proteins and peptides form the chemical foundation of many
378 regulatory processes in the cell. These processes are of direct relevance to the application of
379 metal ions for medicinal[32] and biorthogonal[33] purposes. Xenobiotic metals such as Cd^{2+} target
380 proteins that rely upon nutritive divalent cations for structural and functional integrity. This
381 feature defines the basis for why xenobiotic metals are so deleterious to living cells. From this
382 perspective, characterization of protein interactions with toxic metals in cellular contexts is
383 required for reliable identification of physiologically relevant targets, development of strategies
384 for bioremediation, and creation or refinement of protein-based sensors.

385 In this work, we applied NMR spectroscopy to probe site-specific response of three
386 proteins to Cd^{2+} binding in crowded environments provided by intact cells and their lysates. *E.*
387 *coli* was the experimental model of choice for these studies given its high tolerance to toxic
388 metal ions and the efficiency with which recombinant proteins can be labeled with stable
389 isotopes. Cd^{2+} and the nutritive metals Ca^{2+} and Zn^{2+} were introduced during the window of
390 target protein synthesis and thereby provided a reliable means for generating distinct metalated
391 target protein states in cellular contexts. A major conclusion of our work is that Cd^{2+} targets
392 sulfur-rich, oxygen-rich, as well as mixed sulfur/nitrogen protein coordination sites. That this
393 heavy metal ion successfully competes with native metal ions Ca^{2+} and Zn^{2+} is a surprising result
394 given Cd^{2+} availability is expected to be limited in the highly chelating environment of the
395 cytoplasm. We also demonstrate that Cd^{2+} binding event is associated with distinct protein
396 chemical shift perturbations and modulation of quinary interactions. The distinct spectroscopic
397 signatures associated with metal-ion binding events attest to the potential of NMR spectroscopy
398 as a general tool to probe the interactions of proteins and peptides with toxic metals in crowded
399 biomolecular environments. Moreover, we show that augmentation of NMR experiments with
400 live cell imaging represents a particularly powerful approach that can be extended to other
401 xenobiotic metal ions.
402

Acknowledgements

This work was supported by the National Science Foundation (NSF) grant CHE-1905116 to T.I.I. We acknowledge Dr. Robert Burghardt for assistance with data collection on the Zeiss LSM 780 NLO multiphoton microscope equipped with Airyscan detector (Texas A&M Image Analysis Laboratory, School of Veterinary Medicine & Biomedical Sciences). We thank Ms. Savana M. Green for assistance with handling yeast cultures and collection of fluorescence imaging data, and Mr. Nowlan Savage for his contribution to the development stages of the project. We acknowledge Dr. Vytas A. Bankaitis (Texas A&M University) for guidance on the interpretation of yeast imaging data and for critical reading of the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Author contributions

S.S.K. and T.I.I conceptualized the project and designed the experiments. S.S.K. prepared the samples, collected and analyzed the NMR data, and confocal fluorescence imaging datasets. S.S.K. and T.I.I. wrote the manuscript. T.I.I. supervised the project.

403

References

[1] K. J. Waldron, J. C. Rutherford, D. Ford, N. J. Robinson, *Nature* **2009**, *460*, 823-830.

[2] J. K. Kim, C. Lee, S. W. Lim, A. Adhikari, J. T. Andring, R. McKenna, C.-M. Ghim, C. U. Kim, *Nature Communications* **2020**, *11*, 4557.

[3] K. Krzywoszynska, D. Witkowska, J. Swiatek-Kozlowska, A. Szebesczyk, H. Kozlowski, *Biomolecules* **2020**, *10*.

[4] R. H. Holm, P. Kennepohl, E. I. Solomon, *Chemical Reviews* **1996**, *96*, 2239-2314.

[5] H. Reyes-Caballero, G. C. Campanello, D. P. Giedroc, *Biophys Chem* **2011**, *156*, 103-114.

[6] aJ. Godt, F. Scheidig, C. Grosse-Siestrup, V. Esche, P. Brandenburg, A. Reich, D. A. Groneberg, *J Occup Med Toxicol* **2006**, *1*, 22; bC. C. Bridges, R. K. Zalups, *Toxicol Appl Pharmacol* **2005**, *204*, 274-308.

[7] R. Friedman, *Dalton Transactions* **2014**, *43*, 2878-2887.

[8] aX. Liu, Q. Hu, J. Yang, S. Huang, T. Wei, W. Chen, Y. He, D. Wang, Z. Liu, K. Wang, J. Gan, H. Chen, *Proceedings of the National Academy of Sciences* **2019**, *116*, 20398; bK. A. Morales, Y. Yang, Z. Long, P. Li, A. B. Taylor, P. J. Hart, T. I. Igumenova, *Journal of the American Chemical Society* **2013**, *135*, 12980-12983; cS. Katti, S. B. Nyenhuis, B. Her, A. K. Srivastava, A. B. Taylor, P. J. Hart, D. S. Cafiso, T. I. Igumenova, *Biochemistry* **2017**, *56*, 3283-3295; dT. R. Cole, S. G. Erickson, K. A. Morales, M. Sung, A. Holzenburg, T. I. Igumenova, *Biochemistry* **2019**, *58*, 509-513; eT. R. Cole, T. I. Igumenova, *Frontiers in Molecular Biosciences* **2021**, *8*.

[9] P. Chandrangsu, C. Rensing, J. D. Helmann, *Nature Reviews Microbiology* **2017**, *15*, 338-350.

[10] aL. Barbieri, E. Luchinat, L. Banci, *Nature Protocols* **2016**, *11*, 1101-1111; bD. S. Burz, L. Breindel, A. Shekhtman, in *Methods in Enzymology, Vol. 621* (Ed.: A. K. Shukla), Academic Press, **2019**, pp. 305-328; cE. Luchinat, L. Barbieri, M. Cremonini, A. Nocentini, C. T. Supuran, L. Banci, *Angewandte Chemie International Edition* **2020**, *59*, 6535-6539; dH. Kadavath, N. Cecilia Prymaczok, C. Eichmann, R. Riek, J. A. Gerez, *Angewandte Chemie International Edition* **2023**, *62*, e202213976; eW. Zhu, A. J. Guseman, F. Bhinderwala, M. Lu, X.-C. Su, A. M. Gronenborn, *Angewandte Chemie International Edition* **2022**, *61*, e202201097; fM. Ikari, H. Yagi, T. Kasai, K. Inomata, M. Ito, K. Higuchi, N. Matsuda, Y. Ito, T. Kigawa, *JACS Au* **2023**, *3*, 1658-1669.

[11] L. Banci, L. Barbieri, I. Bertini, E. Luchinat, E. Secci, Y. Zhao, A. R. Aricescu, *Nature Chemical Biology* **2013**, *9*, 297-299.

[12] Y. Ye, X. Liu, G. Xu, M. Liu, C. Li, *Angewandte Chemie International Edition* **2015**, *54*, 5328-5330.

[13] aD. H. Nies, *Appl Microbiol Biotechnol* **1999**, *51*, 730-750; bW. Qin, J. Zhao, X. Yu, X. Liu, X. Chu, J. Tian, N. Wu, *Frontiers in Microbiology* **2019**, *10*.

[14] aC. Block, S. Freyermuth, D. Beyersmann, A. N. Malviya, *Journal of Biological Chemistry* **1992**, *267*, 19824-19828; bG. J. Long, *Toxicology and Applied Pharmacology* **1997**, *143*, 189-195.

[15] aG. Zhang, M. G. Kazanietz, P. M. Blumberg, J. H. Hurley, *Cell* **1995**, *81*, 917-924; bS. S. Katti, I. V. Krieger, J. Ann, J. Lee, J. C. Sacchettini, T. I. Igumenova, *Nature Communications* **2022**, *13*, 2695.

[16] aM. Guerrero-Valero, C. Ferrer-Orta, J. Querol-Audí, C. Marin-Vicente, I. Fita, J. C. Gómez-Fernández, N. Verdaguer, S. Corbalán-García, *Proceedings of the National Academy of Sciences* **2009**, *106*, 6603; bA. Honigmann, G. van den Bogaart, E. Iraheta, H. J. Risselada, D. Milovanovic, V. Mueller, S. Müllar, U. Diederichsen, D. Fasshauer, H. Grubmüller, S. W. Hell, C. Eggeling, K. Kühnel, R. Jahn, *Nature Structural & Molecular Biology* **2013**, *20*, 679-686.

[17] S. Chakraborty, J. Yudenfreund Kravitz, P. W. Thulstrup, L. Hemmingsen, W. F. DeGrado, V. L. Pecoraro, *Angewandte Chemie International Edition* **2011**, *50*, 2049-2053.

[18] J. S. Plegaria, S. P. Dzul, E. R. Zuiderweg, T. L. Stemmler, V. L. Pecoraro, *Biochemistry* **2015**, *54*, 2858-2873.

[19] aX. Mu, S. Choi, L. Lang, D. Mowray, N. V. Dokholyan, J. Danielsson, M. Oliveberg, *Proceedings of the National Academy of Sciences* **2017**, *114*, E4556; bS. Leeb, T. Sörensen, F. Yang, X. Mu, M. Oliveberg, J. Danielsson, *Current Research in Structural Biology* **2020**, *2*, 68-78.

[20] D. S. Burz, L. Breindel, A. Shekhtman, *Int J Mol Sci* **2019**, *20*.

[21] K. A. Morales, T. I. Igumenova, *Biochemistry* **2012**, *51*, 3349-3360.

[22] M. D. Stewart, T. I. Igumenova, *Biochemistry* **2012**, *51*, 7263-7277.

[23] Y. Xu, L. Feng, P. D. Jeffrey, Y. Shi, F. M. M. Morel, *Nature* **2008**, *452*, 56-61.

[24] D. G. Nettesheim, H. R. Engeseth, J. D. Otvos, *Biochemistry* **1985**, *24*, 6744-6751.

[25] aM. Huang, I. C. Shaw, D. H. Petering, *J Inorg Biochem* **2004**, *98*, 639-648; bR. Kothinti, A. Blodgett, N. M. Tabatabai, D. H. Petering, *Chemical Research in Toxicology* **2010**, *23*, 405-412; cI. V. Smirnova, D. C. Bittel, R. Ravindra, H. Jiang, G. K. Andrews, *J Biol Chem* **2000**, *275*, 9377-9384.

[26] M. A. Namdarghanbari, J. Bertling, S. Krezoski, D. H. Petering, *J Inorg Biochem* **2014**, *136*, 115-121.

[27] M. Borsari, in *Encyclopedia of Inorganic and Bioinorganic Chemistry*, **2014**, pp. 1-16.

[28] D. Wang, S. Huang, P. Liu, X. Liu, Y. He, W. Chen, Q. Hu, T. Wei, J. Gan, J. Ma, H. Chen, *Scientific Reports* **2016**, *6*, 33391.

[29] J. Kim, D. J. Klionsky, *Annual Review of Biochemistry* **2000**, *69*, 303-342.

[30] aF. Yu, V. M. Cangelosi, M. L. Zastrow, M. Tegoni, J. S. Plegaria, A. G. Tebo, C. S. Mocny, L. Ruckthong, H. Qayyum, V. L. Pecoraro, *Chem Rev* **2014**, *114*, 3495-3578; bJ. S. Plegaria, V. L. Pecoraro, *Methods Mol Biol* **2016**, *1414*, 187-196.

[31] aP. M. S. Bragança, M. S. P. Carepo, S. R. Pauleta, T. B. J. Pinter, M. Elia, C. M. Cordas, I. Moura, V. L. Pecoraro, J. J. G. Moura, *Journal of Inorganic Biochemistry* **2023**, *240*, 112096; bK. J. Koebke, T. B. J. Pinter, W. C. Pitts, V. L. Pecoraro, *Chemical Reviews* **2022**, *122*, 12046-12109.

[32] aC. Di Natale, I. De Benedictis, A. De Benedictis, D. Marasco, *Antibiotics (Basel)* **2020**, *9*; bE. Boros, P. J. Dyson, G. Gasser, *Chem* **2020**, *6*, 41-60.

[33] M. O. N. van de L'Isle, M. C. Ortega-Liebana, A. Unciti-Broceta, *Current Opinion in Chemical Biology* **2021**, *61*, 32-42.