Supplementary Information

to

A histidine-rich *Pseudomonas* metallothionein with a disordered tail displays higher binding capacity for cadmium than zinc

Jelena Habjanič, Oliver Zerbe, Eva Freisinger

Table of contents

Figure S1. Amino acid sequence alignment of MTs from different *Pseudomonas* species and strains. Cys residues are highlighted in black, His residues in grey. 1) *P. fluorescens* strain Pf0-1 (UniProt[1] accession number: Q3K9Q2), 2) *P. fluorescens* R124 (K0X1F5), 3) *P. fluorescens* SBW25 (C3K8N4), 4) *P. fluorescens* BRIP34879 (L7HBL7), 5) *P. fluorescens* BBc6R8 (V7E9L0), 6) *P. aeruginosa* PAO1 (Q9I1X5), 7) *P. putida* KT2440 (Q88HU1), 8) *P. fluorescens* Q2-87 (J2EKT7), 9) *P. fluorescens* BTR19 (A0A1Q5X1W4), 10) *S. elongatus* PCC 7942 (P30331).

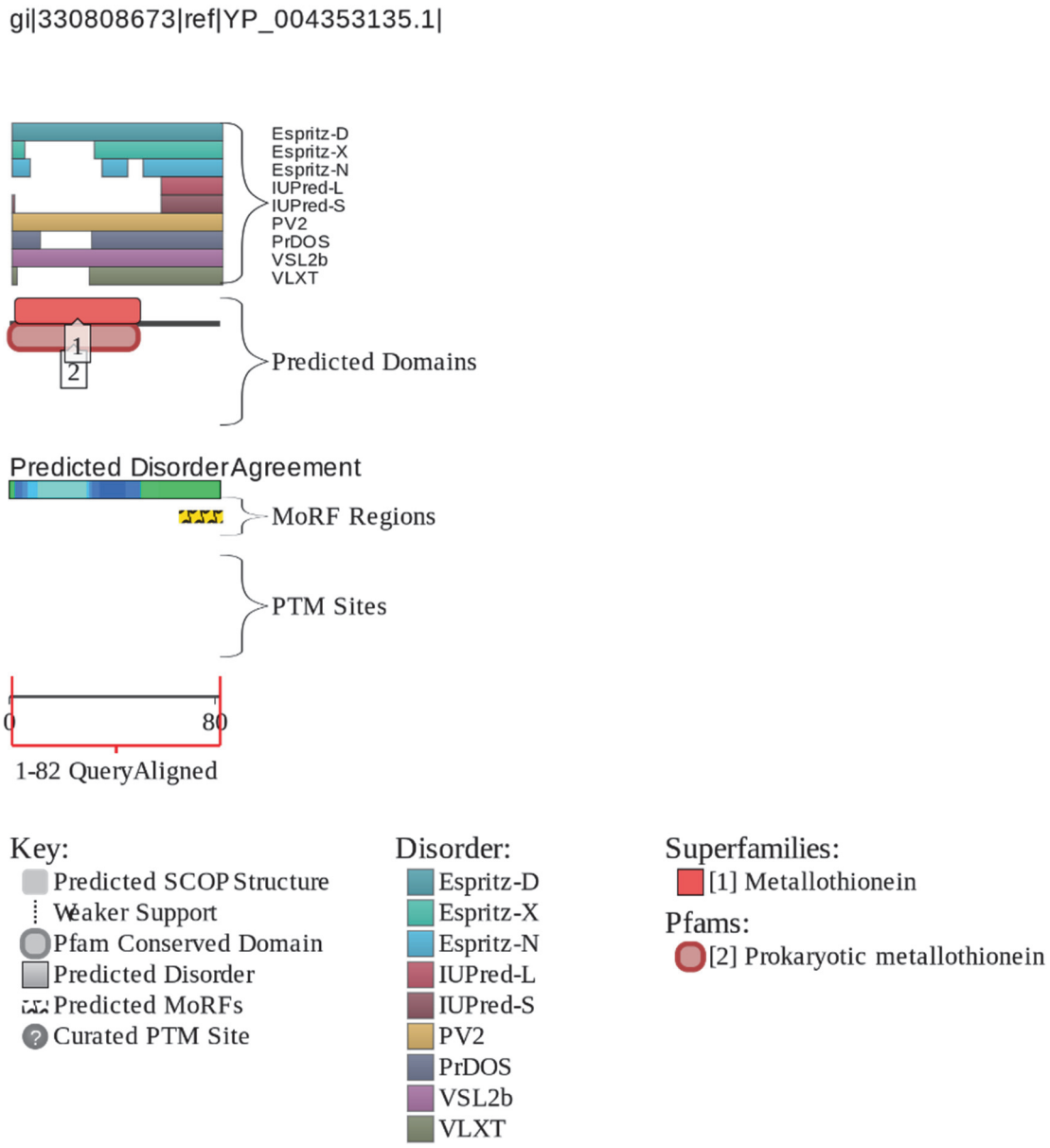

Figure S2. Graphical representation of results obtained from the D^2P^2 (Database of Disordered Protein Predictions)[2] for PflQ2 MT.

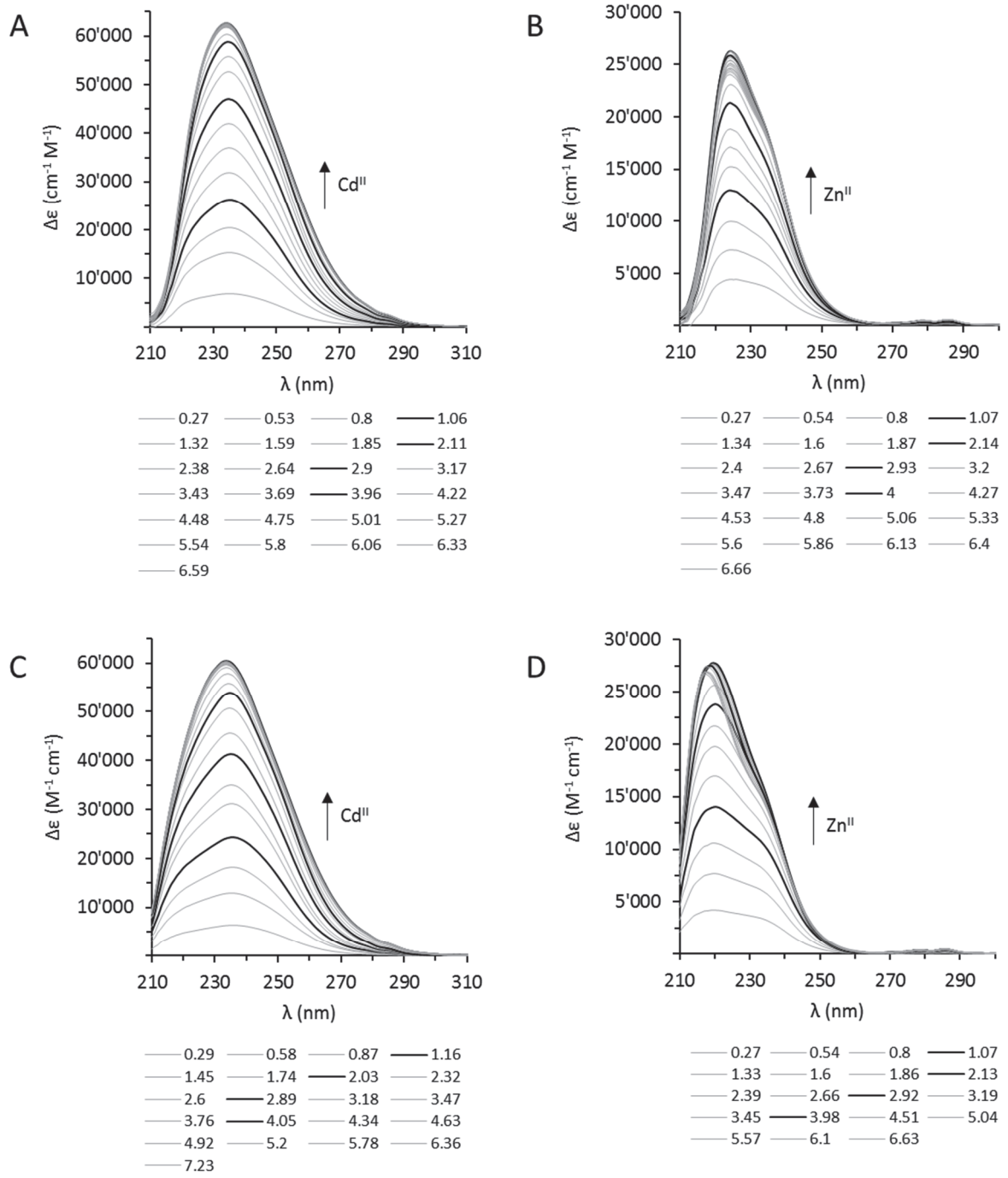

Figure S3. UV spectra of titration of the apo-form of PflQ2 MT (A, B) and sh_PflQ2 MT (C, D) with the indicated metal ions.

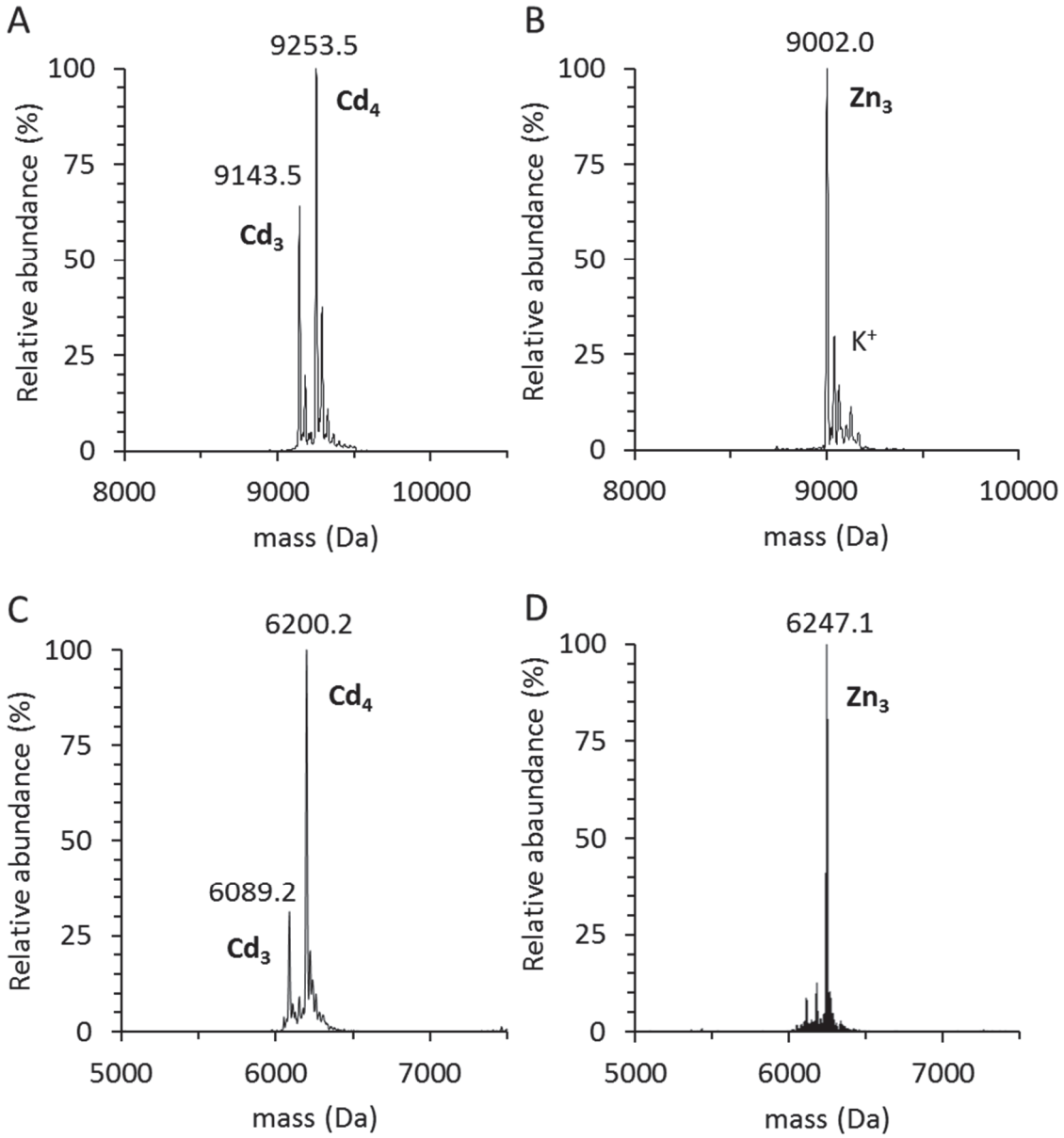

Figure S4. Deconvoluted ESI(+) mass spectra for Zn^{II} and Cd^{II} species of PflQ2 MT (A: M_{calc} 9253.20 Da (Cd_4), 9140.7 9 Da (Cd_3); B: M_{calc} 8999.7 Da (Zn_3)) and sh_PflQ2 MT (C: M_{calc} 6199.92 Da (Cd_4), 6087.51 Da (Cd_3), D: M_{calc} 6250.16 Da (Zn_3, ^{13}C, ^{15}N labelled protein)).

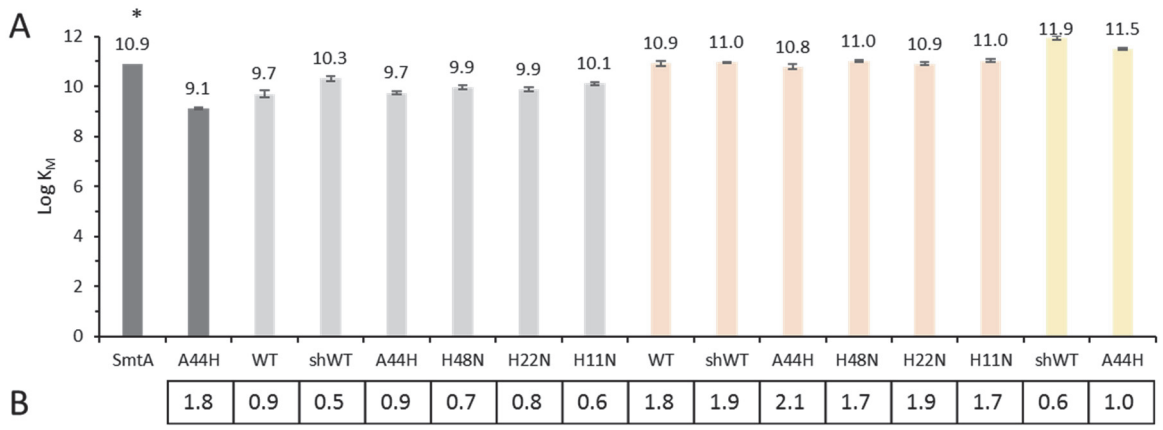

Figure S5. A) Log values for average apparent binding constants (I=321 mM) for Zn_4 (dark grey), Zn_3 (light grey), Cd_4 (red), Cd_3 (orange) species of PflQ2 MT (WT) and sh_PflQ2 MT (shWT) as well as its mutants. The reference value for Zn_4SmtA (I=93 mM) is taken from the literature.[3] B) Amount of metal ion equivalents removed by 5F-BAPTA during the experiment.

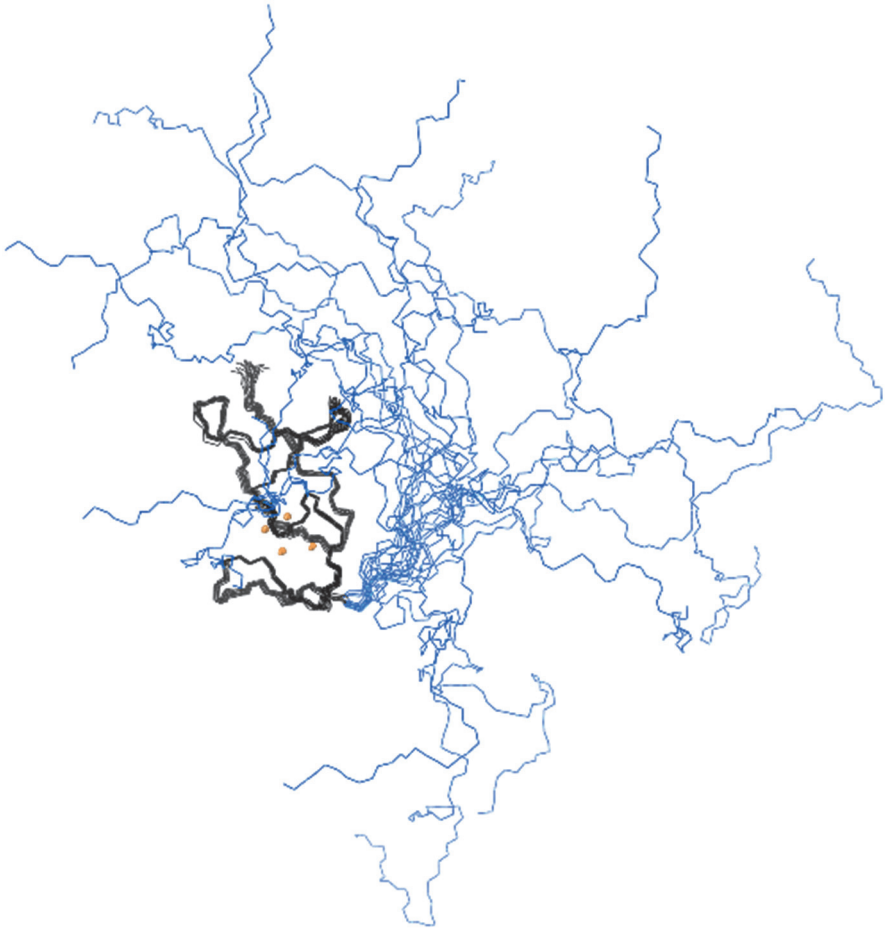

Figure S6. Ensemble of the 20 lowest-energy structures of Cd_4PflQ2 MT. Residues 51-81 of the C-terminal tail are coloured in blue.

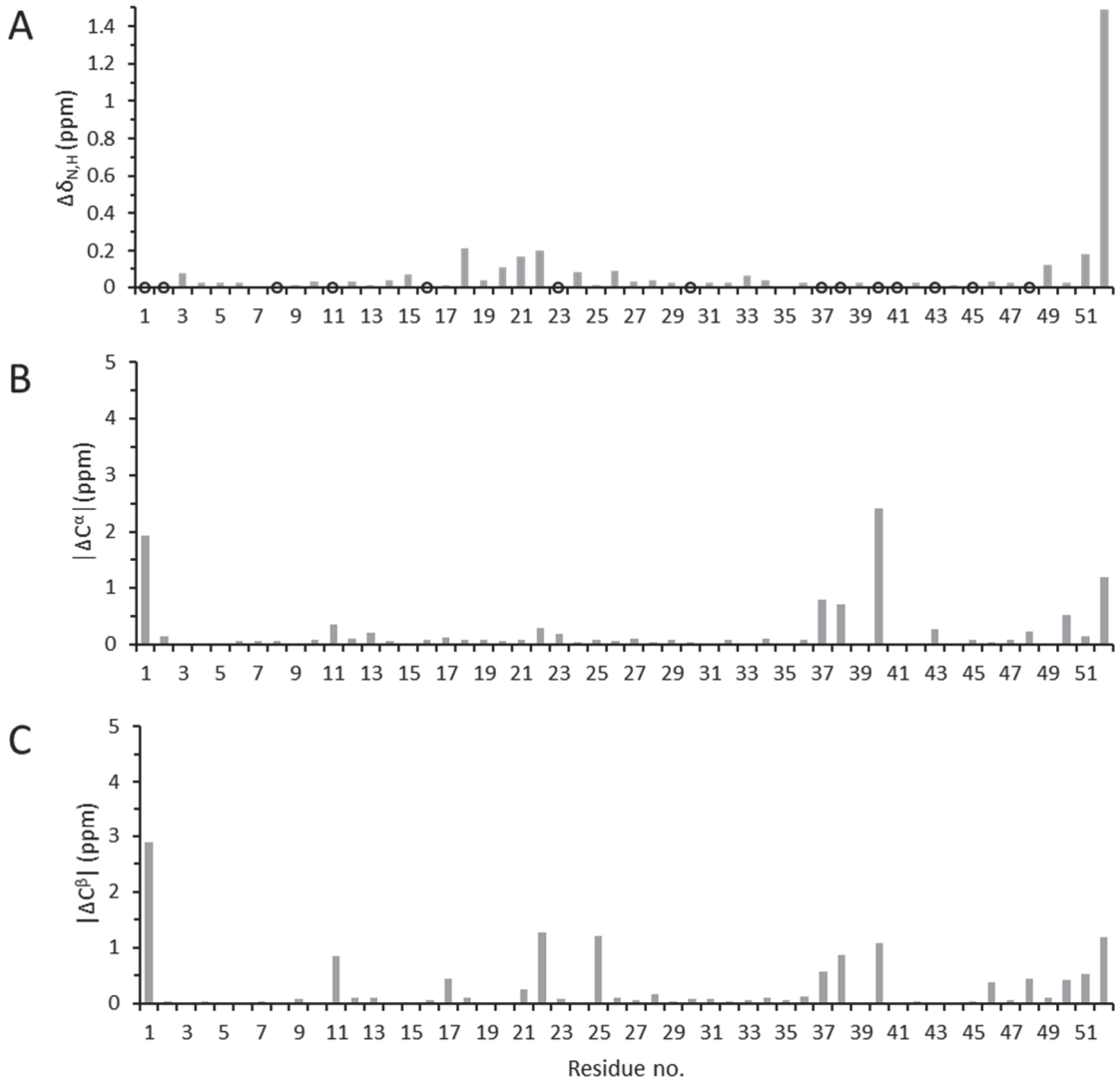

Figure S7. Analysis of chemical shift differences of Cd_4 species of PflQ2 MT and sh_PflQ2 MT for A) backbone, B) C^{α}, and C) C^{β}.

.

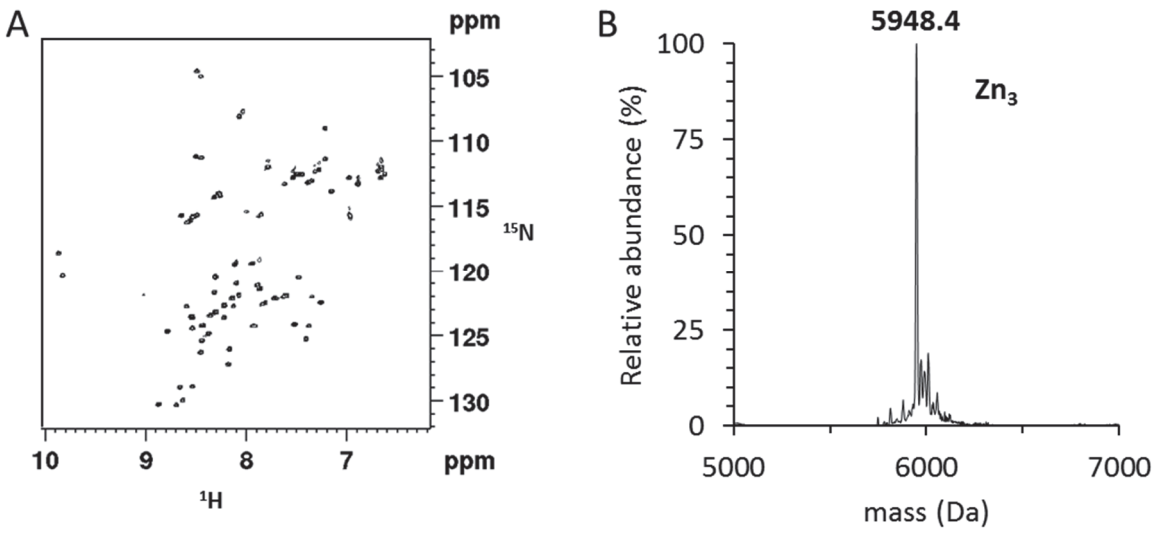

Figure S8. Analysis of "Zn_4" species of sh_PflQ2 MT: A) [^{15}N,^{1}H]-HSQC spectrum, B) Deconvoluted ESI(+) mass spectrum (M_{calc} 5946.42 Da (Zn_3), 6011.8 Da (Zn_4)).

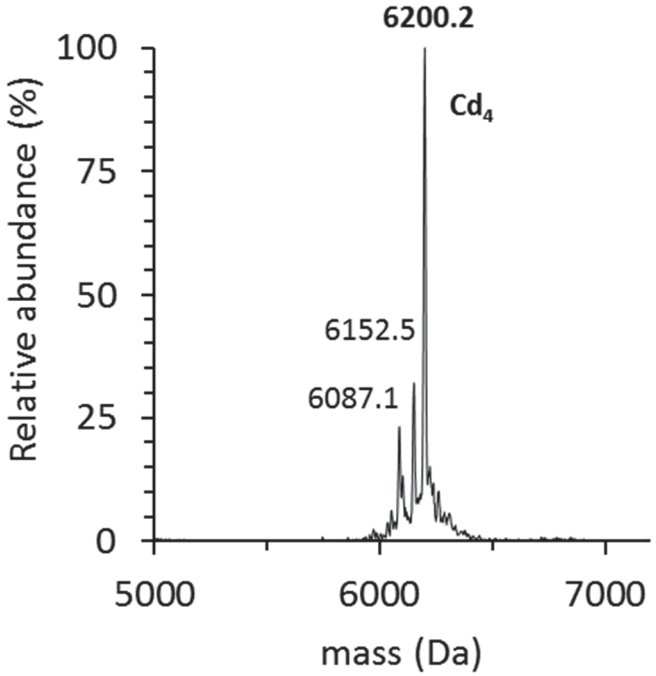

Figure S9. Deconvoluted ESI(+) mass spectrum of Zn^{II}-to-Cd^{II} exchange of sh_PflQ2 MT (M_{calc} 6199.92 Da (Cd_4)).

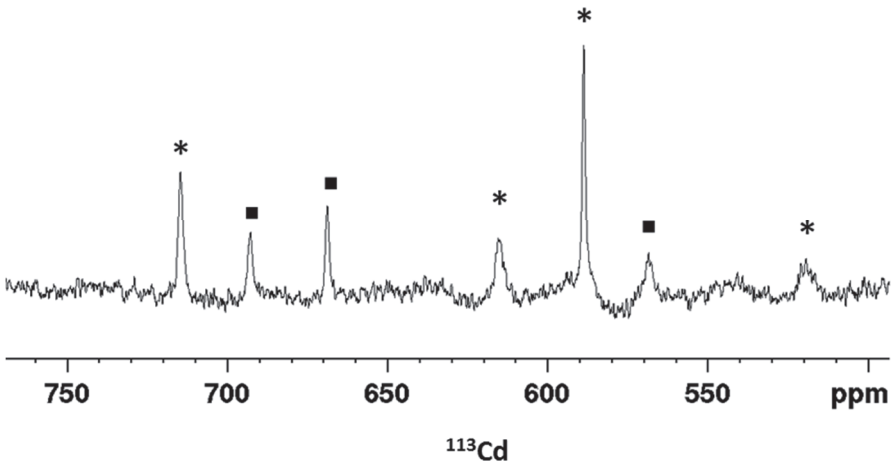

Figure S10. Proton-decoupled 1D ^{113}Cd spectrum at 320 K of a sample containing a mixture of $^{113}Cd_4$-(*) and $^{113}Cd_3$sh_PflQ2 MT (▪).

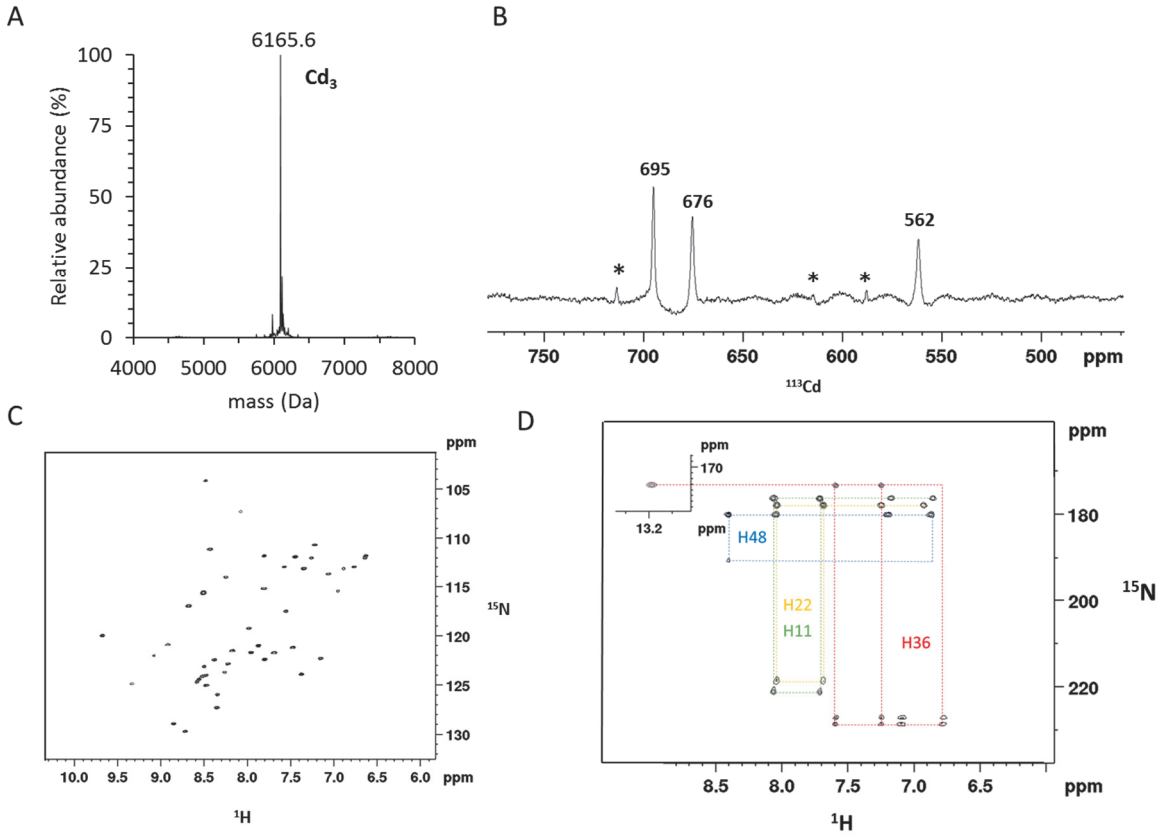

Figure S11. A) Deconvoluted ESI(+) mass spectrum (M_{calc}.6161.95 Da (Cd_3, ^{15}N labelled protein) , B) 1D ^{113}Cd spectrum (320 K, note that due to high concentrations needed for the measurement of ^{113}Cd spectra, background formation of additional cadmium peaks with chemical shifts corresponding to Cd_4 species are always observed), C) [^{15}N,^{1}H]-HSQC spectrum, and D) long-range [^{15}N,^{1}H]-HSQC spectrum (J(^{15}N, ^{1}H) =22.75 Hz; 310 K) of ^{13}C, ^{15}N labelled $^{113}Cd_3$sh_PflQ2 MT.

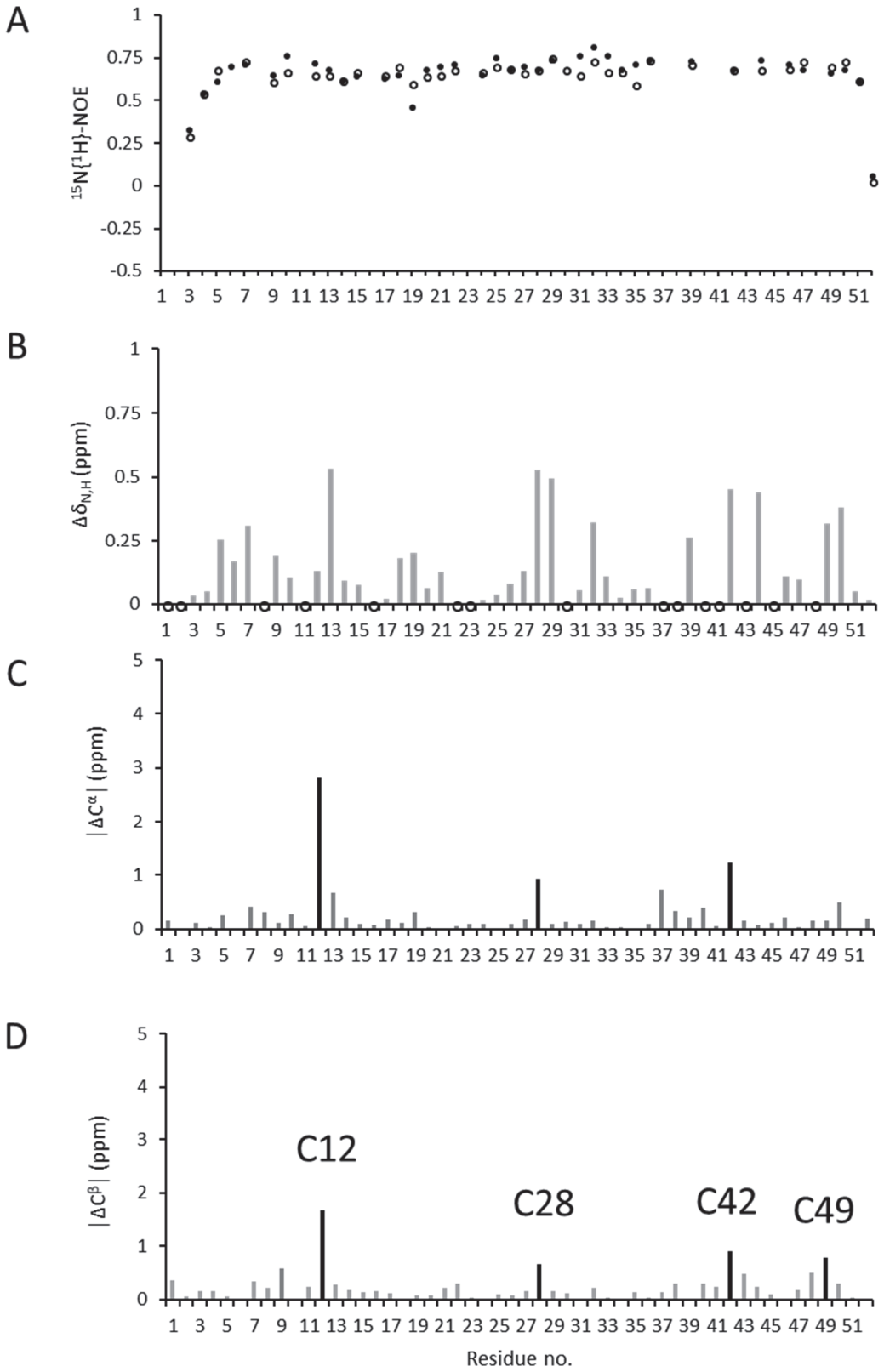

Figure S12. Comparison of backbone dynamics and chemical shift differences between Cd_4- and Cd_3sh_PflQ2 MT: A) $^{15}N\{^1H\}$-NOE values for Cd_4 (full circles) and Cd_3 (empty circles). B) Backbone chemical shift perturbation (values that could not be determined are indicated by o). Chemical shift differences for C^{α} (C) and C^{β} (D).

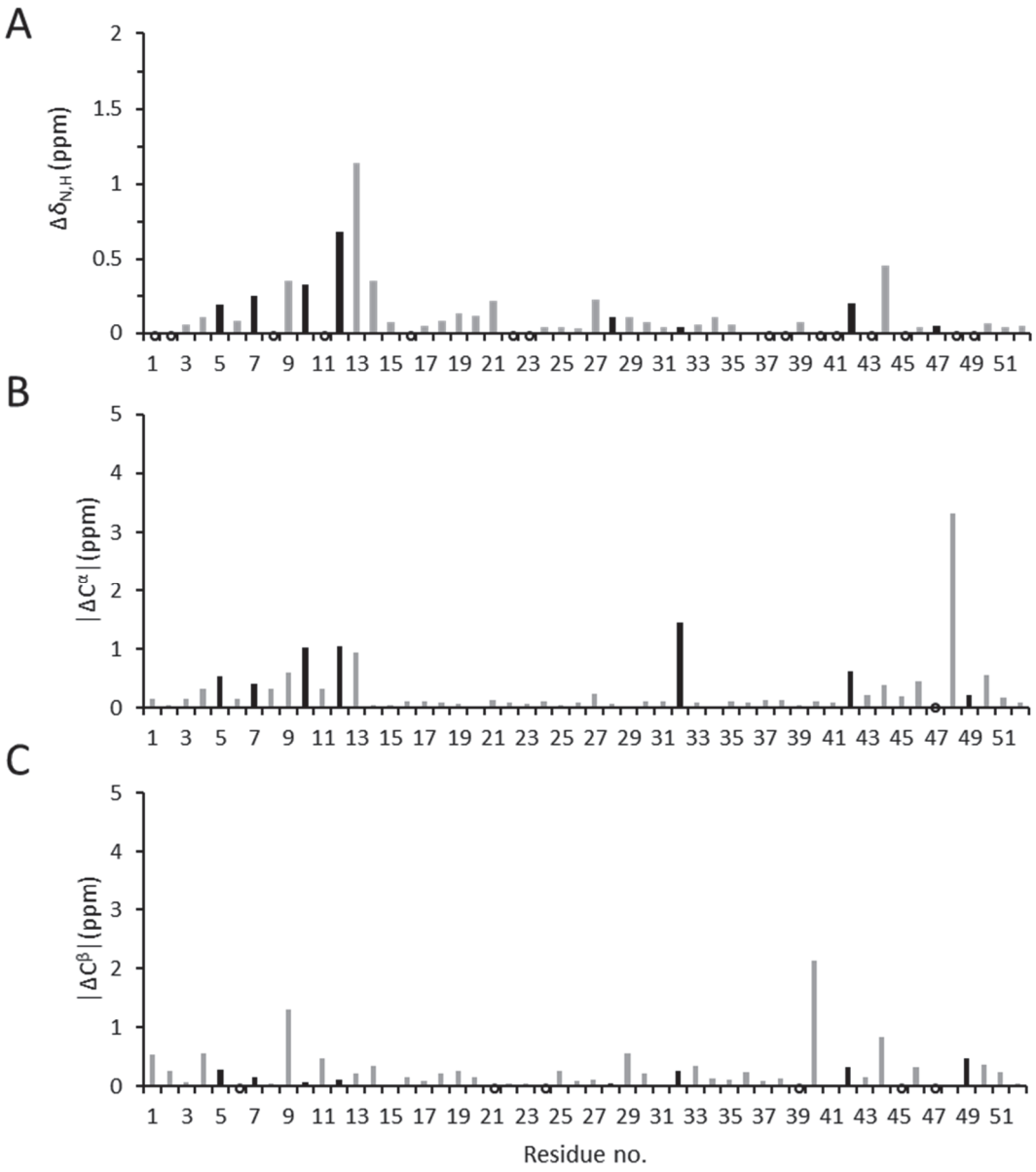

Figure S13. Chemical shift perturbation differences between Zn_3- and Cd_3sh_PflQ2 MT for the backbone (A), C^{α} (B) and C^{β} (C). Cysteine residues are shown with black bars while data points that could not be derived unambiguously are indicated by o.

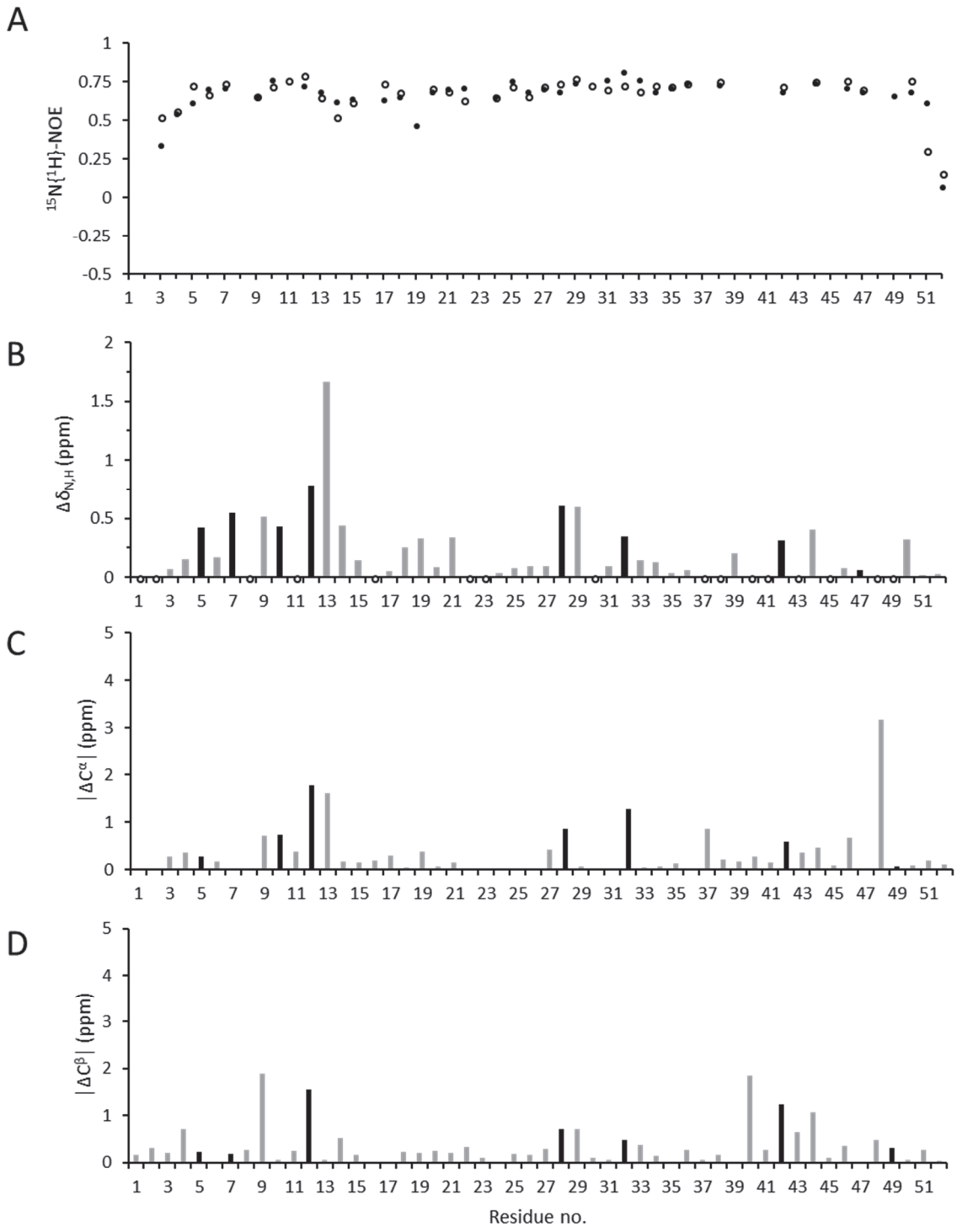

Figure S14. Comparison of backbone dynamics and chemical shift differences for Cd_4- and Zn_3sh_PflQ2 MT: A) $^{15}N\{^1H\}$-NOE values for Cd_4 (full circles) and Zn_3 (empty circles). B-D) Chemical shift changes for the backbone (B), C^{α} (C), and C^{β} (D).

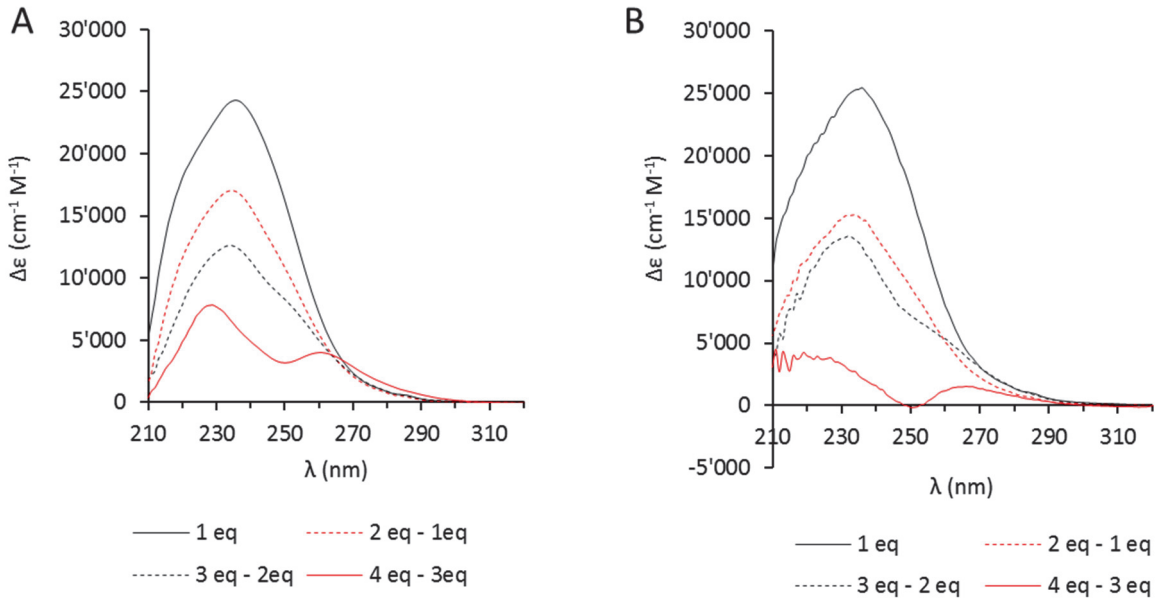

Figure S15. Difference UV spectra of the titration of A) apo-sh_PflQ2 MT (see Figure S6C for the original spectra) and B) apo-A44H-sh_PflQ2 MT with Cd^{II} ions. In both cases, a bathochromic shift of the absorption (around 260-265 nm) that is indicative of Cd^{II}-thiolate cluster formation is observed upon addition of the 3rd and of the 4th equivalent of Cd^{II}.

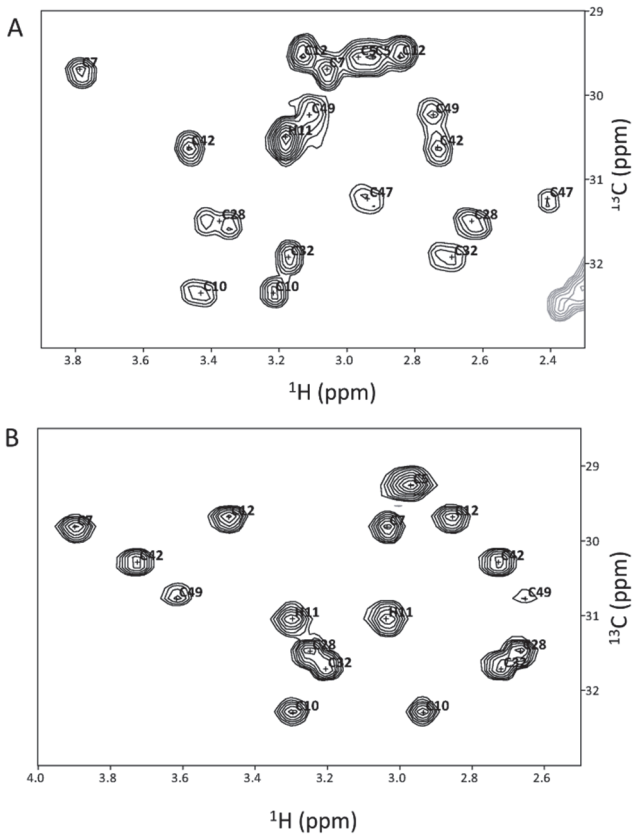

Figure S16. Assignment of C^{β}/H^{β} correlations of Cys residues in [^{13}C,^{1}H]-HSQC spectra for A) Cd_3- and B) Zn_3sh_PflQ2 MT. For the Zn_3 species, C47 was not observed in the [^{13}C,^{1}H]-HSQC spectrum.

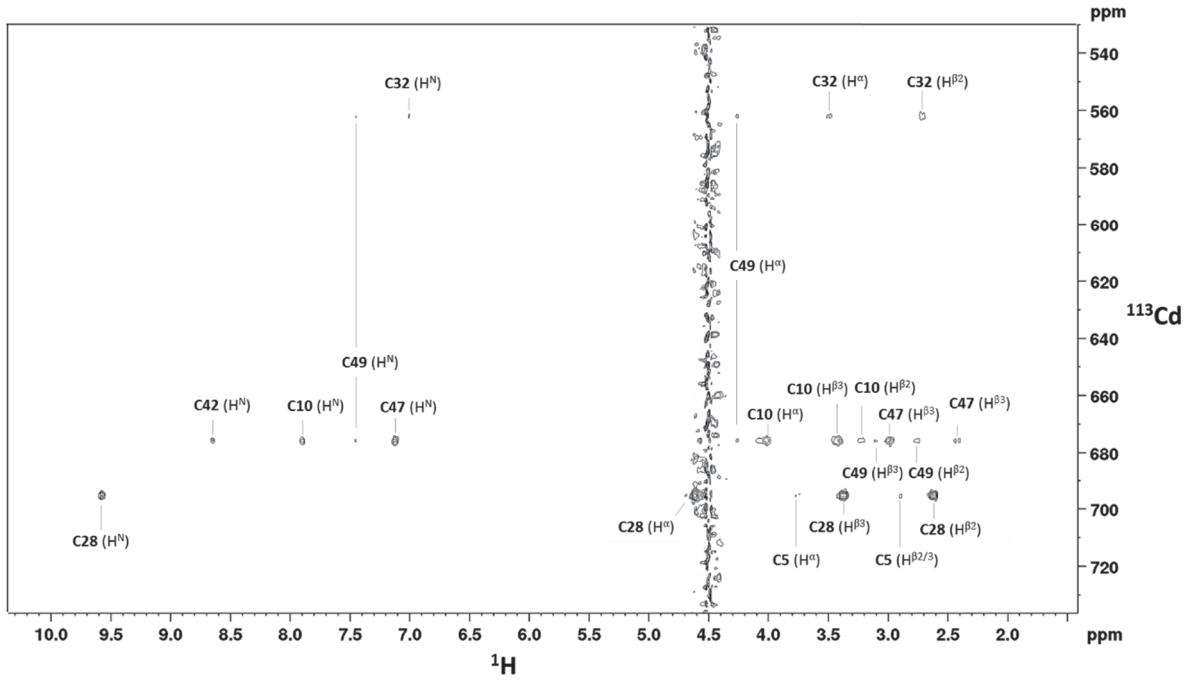

Figure S17. [^{113}Cd,^{1}H]-HSQC-TOCSY of $^{113}Cd_3$sh_PflQ2 MT (320 K, J(^{113}Cd ^{1}H)=35 Hz).

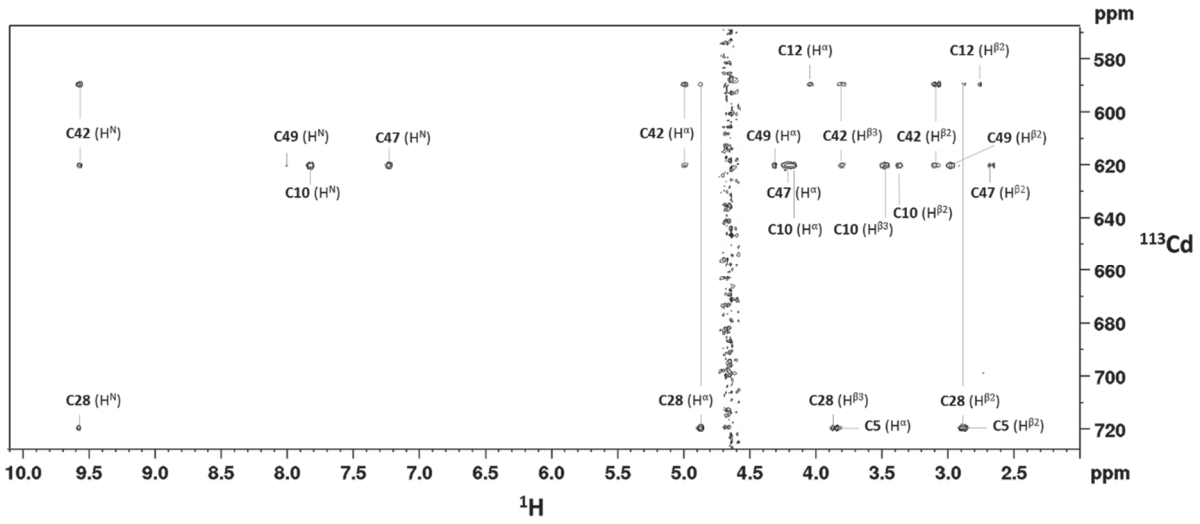

Figure S18. [^{113}Cd,^{1}H]-HSQC-TOCSY of $^{113}Cd_4$sh_PflQ2 MT (310 K, J(^{113}Cd ^{1}H)=40 Hz).

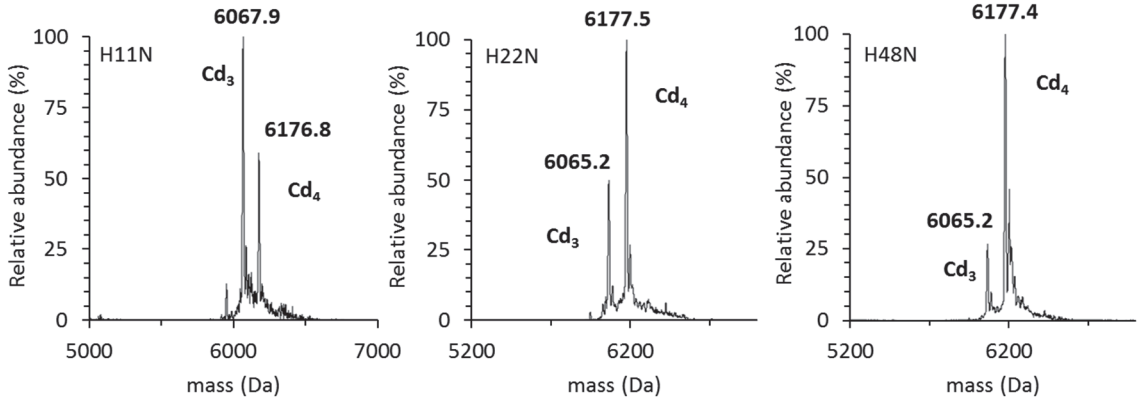

Figure S19. Deconvolved ESI(+) mass spectra of Cd_4 species of His-to-Asn mutants (M_{calc}. 6064.24 Da (Cd_3), 6177.04 Da (Cd_4)).

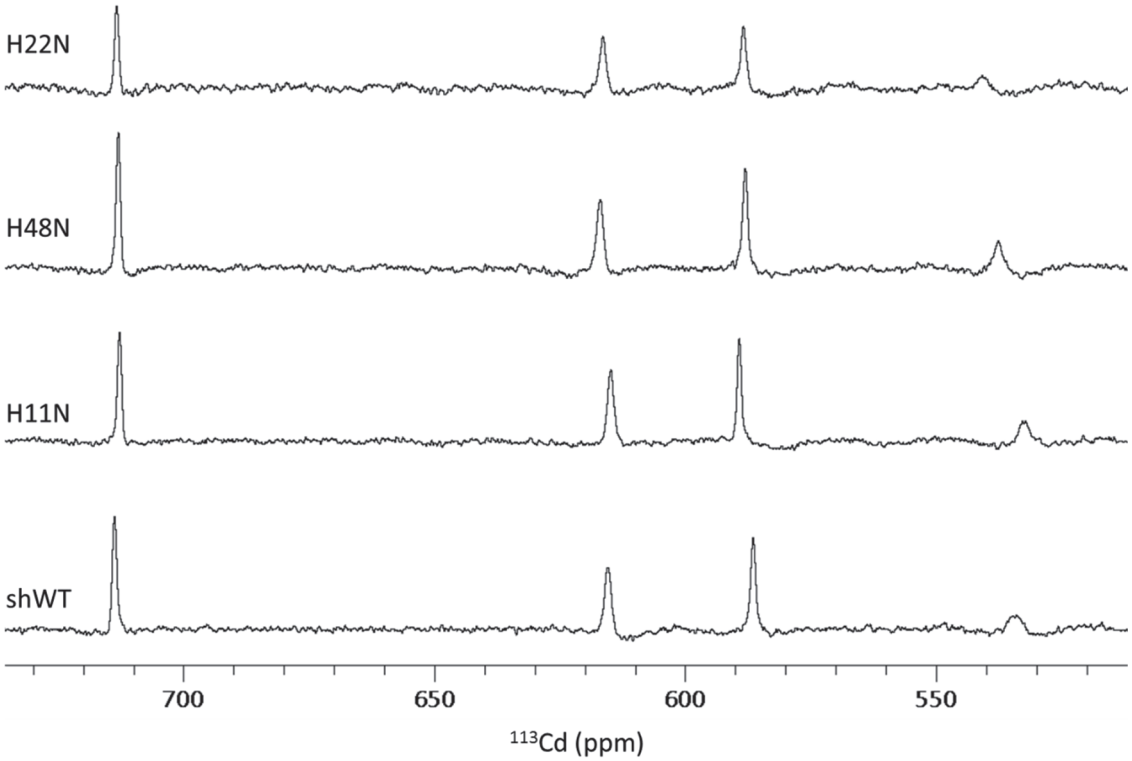

Figure S20. Comparison of the proton-decoupled 1D ^{113}Cd spectra of Cd_4 species of sh_PflQ2 MT (shWT) and its His-to-Asn mutants at 320 K.

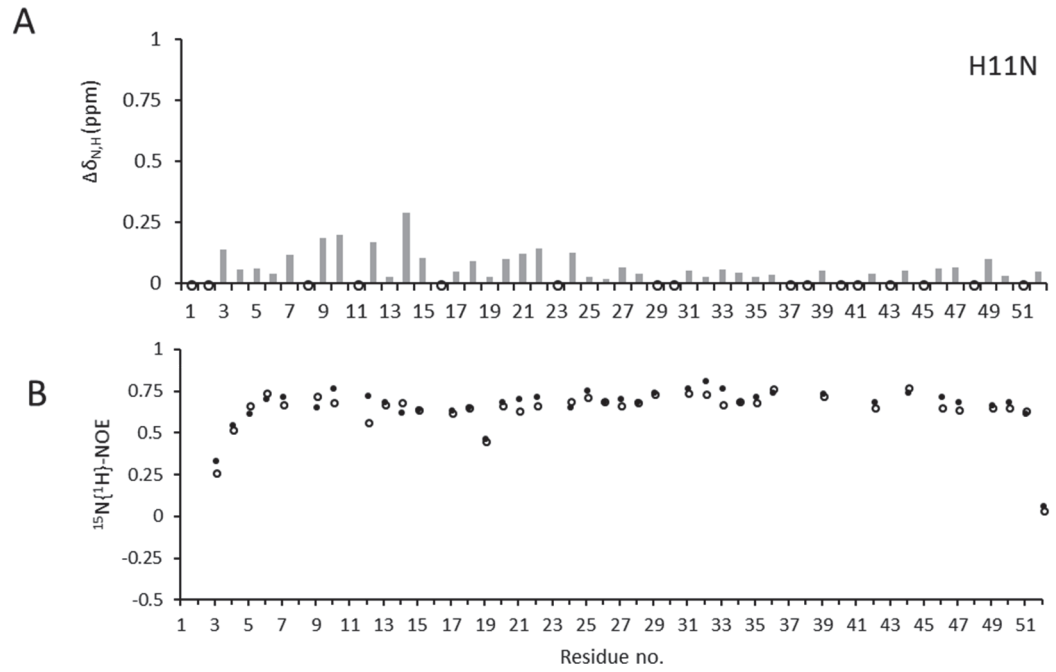

Figure S21. Comparison of Cd_4 species of sh_PflQ2 MT and its H11N mutant: A) Backbone chemical shift changes (o symbol depicts values that could not be determined) B) ^{15}N{^{1}H}-NOE values for sh_PflQ2 MT (full circles) and its H11N mutant (empty circles).

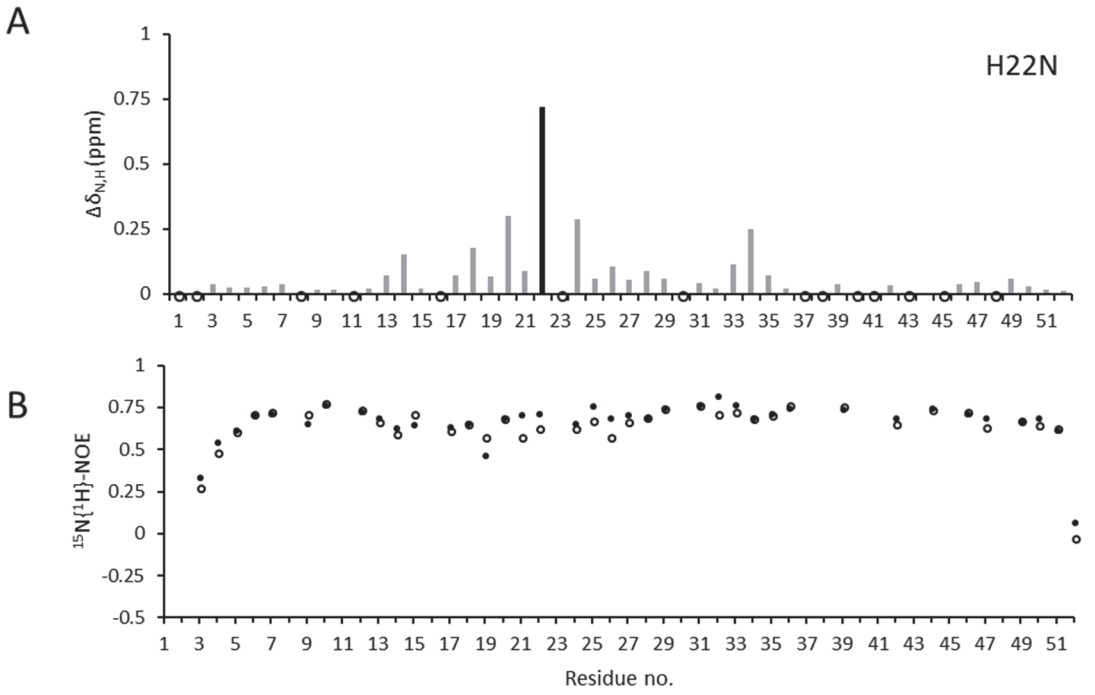

Figure S22. Comparison of Cd_4 species of sh_PflQ2 MT and its H22N mutant: A) Backbone chemical shift changes (o symbol depicts values that could not be determined) B) $^{15}N\{^{1}H\}$-NOE values for sh_PflQ2 MT (full circles) and its H22N mutant (empty circles).

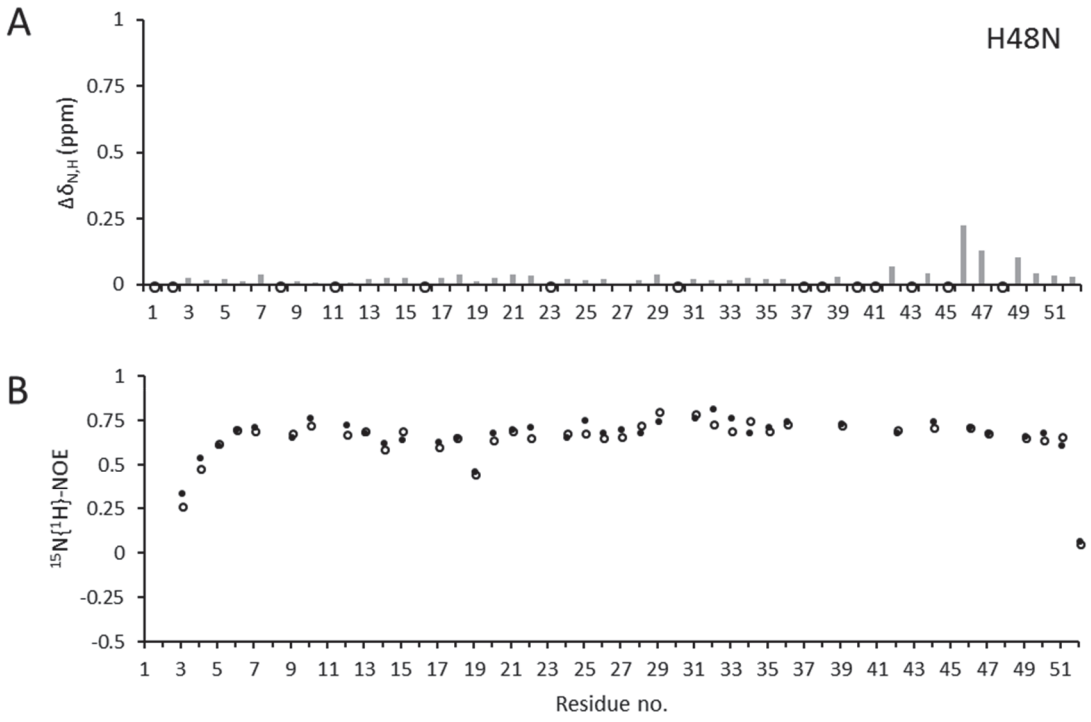

Figure S23. Comparison of Cd_4 species of sh_PflQ2 MT and its H48N mutant: A) Backbone chemical shift differences (values that could not be determined are indicated by **o**) B) $^{15}N\{^{1}H\}$-NOE values for sh_PflQ2 MT (full circles) and its H48N mutant (empty circles).

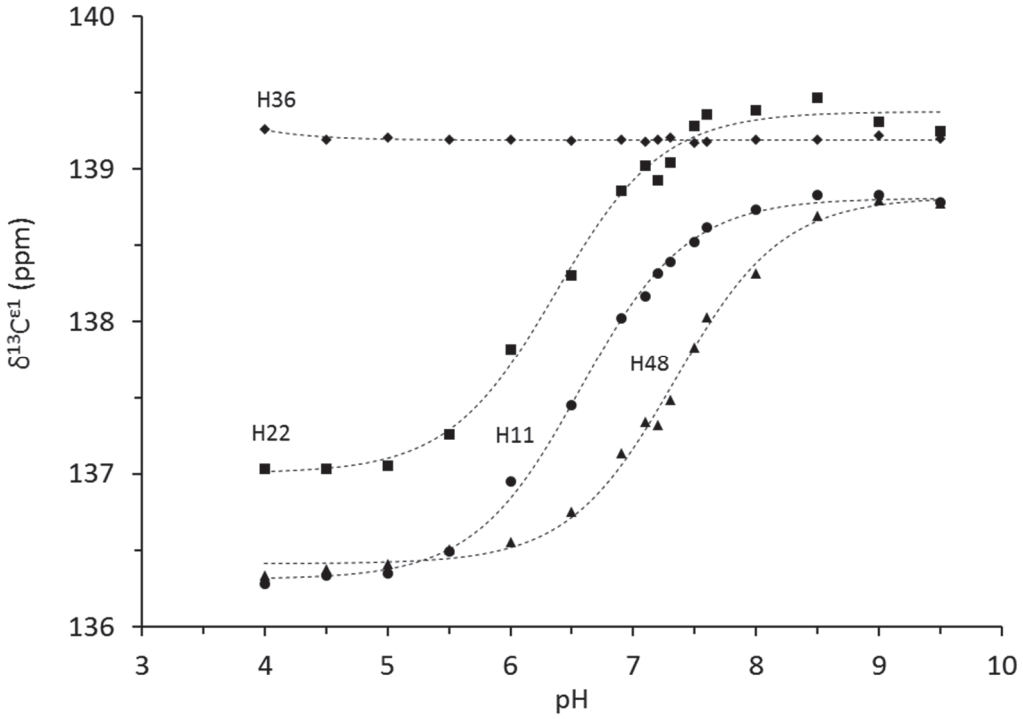

Figure S24. Plot of $C^{\varepsilon 1}$ chemical shifts for the four histidine residues in ^{15}N, ^{13}C labelled Cd_4sh_PflQ2 MT with decreasing pH.

Table S1. Fitting data for determination of pK_a values of histidine residues for Cd_4sh_PflQ2 MT.

Equation $$\delta_{obs} = \frac{\delta_{HA} + \delta_A \cdot 10^{pH-pK_a}}{1 + 10^{pH-pK_a}}$$

	Residue	**H11**	**H22**	**H48**
$^{1}H^{\varepsilon 1}$	δ_{HA}	7.63476 ± 0.00537	7.68905 ± 0.00929	7.65545 ± 0.02164
	δ_A	8.54295 ± 0.00606	8.58688 ± 0.01146	8.53726 ± 0.01534
	pK_1	6.58873 ± 0.01691	6.46612 ± 0.03253	7.39965 ± 0.04444
	Reduced Chi-Sqr	1.33E-04	4.41E-04	0.00123
	R-Square(COD)	0.99913	0.99706	0.9901
	Adj. R-Square	0.99899	0.9966	0.98858
$^{13}C^{\varepsilon 1}$	δ_{HA}	138.81136 ± 0.02043	139.37344 ± 0.03692	138.81622 ± 0.03801
	δ_A	136.30922 ± 0.0233	137.00155 ± 0.04901	136.41352 ± 0.02778
	pK_1	6.57362 ± 0.02362	6.3706 ± 0.05252	7.34241 ± 0.02897
	Reduced Chi-Sqr	0.00195	0.00756	0.00395
	R-Square(COD)	0.99832	0.99277	0.99576
	Adj. R-Square	0.99806	0.99165	0.99511
	Average pK_1	**6.58**	**6.42**	**7.37**

Figure S25. Structure comparison of Zn_4SmtA (red backbone, Zn^{II} are represented as grey spheres)[3] and Cd_4sh_PflQ2MT (black backbone, Cd^{II} are represented as orange spheres).

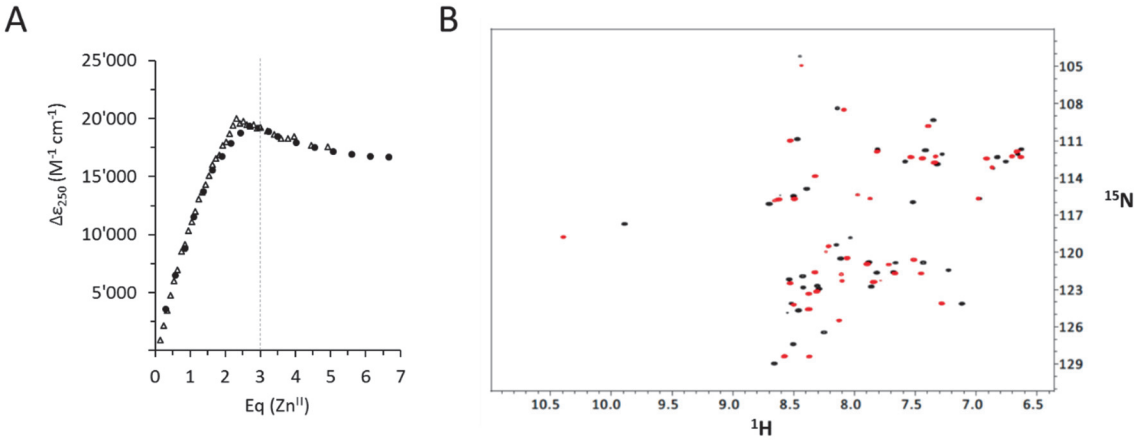

Figure S26. A) Plot of thiolate-to-Zn^{II} charge transfer band at 230 nm for the titration of apo-sh_PflQ2 (●) and apo-A44H-sh_PflQ2 MT (Δ) with Zn^{II} ions; B) Overlay of [^{15}N,^{1}H]-HSQC spectra of Cd_4A44H-sh_PflQ2 (black) and Zn_4A44H-sh_PflQ2 MT (red).

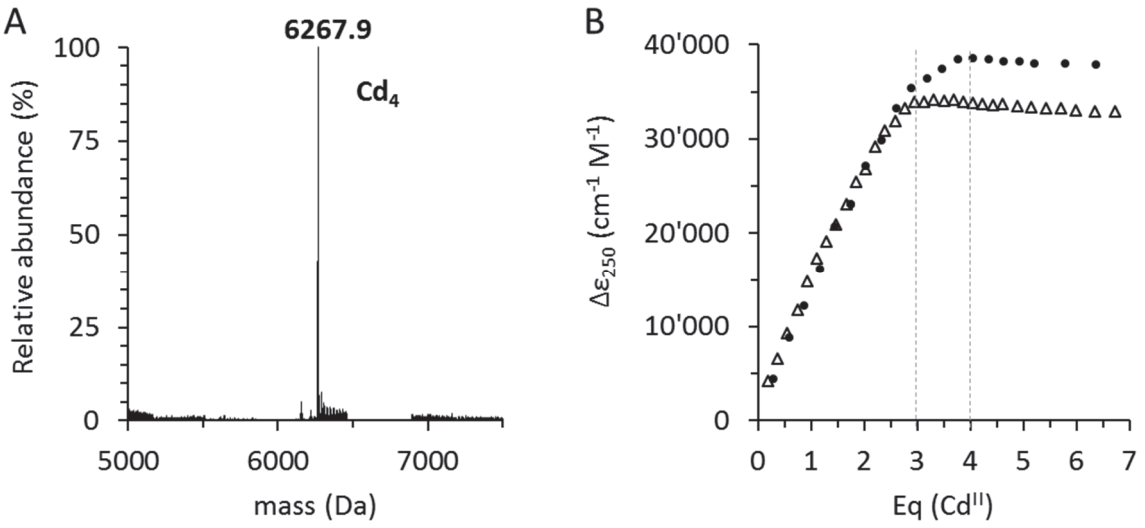

Figure S27. A) Deconvolved ESI(+) mass spectrum of Cd_4A44H-sh_PflQ2 MT (M_{calc} 6265.98 Da (Cd_4)). B) The plot of thiolate-to-Cd^{II} charge transfer band at 250 nm for the titration of apo-sh_PflQ2 (●) and apo-A44H_sh_PflQ2 MT (Δ) with Cd^{II} ions.

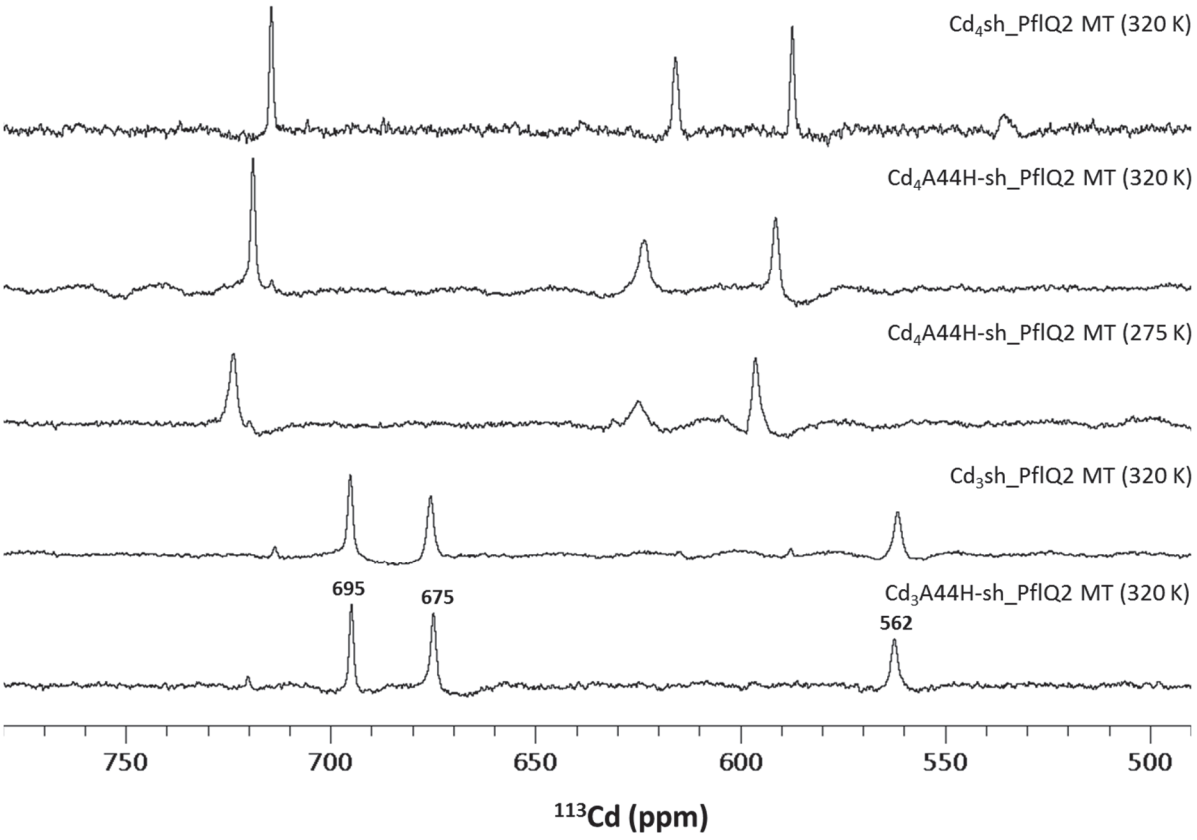

Figure S28. Comparison of proton-decoupled 1D ^{113}Cd spectra of different Cd^{II} species of sh_PflQ2 MT and its A44H mutant.

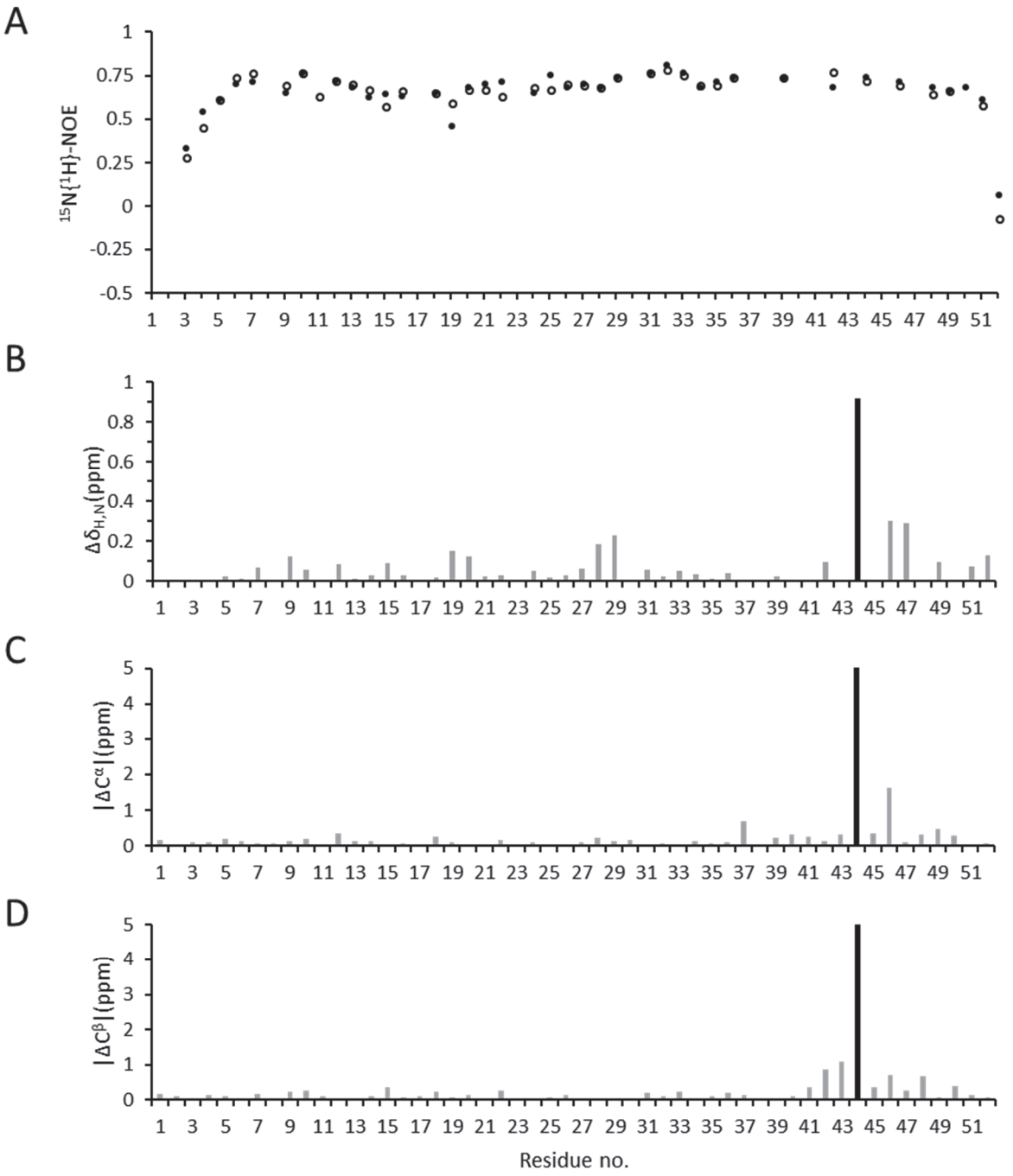

Figure S29. Comparison of Cd_4 species of sh_PflQ2 MT and its A44H mutant: A) $^{15}N\{^{1}H\}$-NOE values for sh_PflQ2 MT (full circles) and its A44H mutant (empty circles). B-D) Chemical shift changes for the backbone (B), C^{α} (C), and C^{β} (D) (values that could not be determined are indicated by **o**).

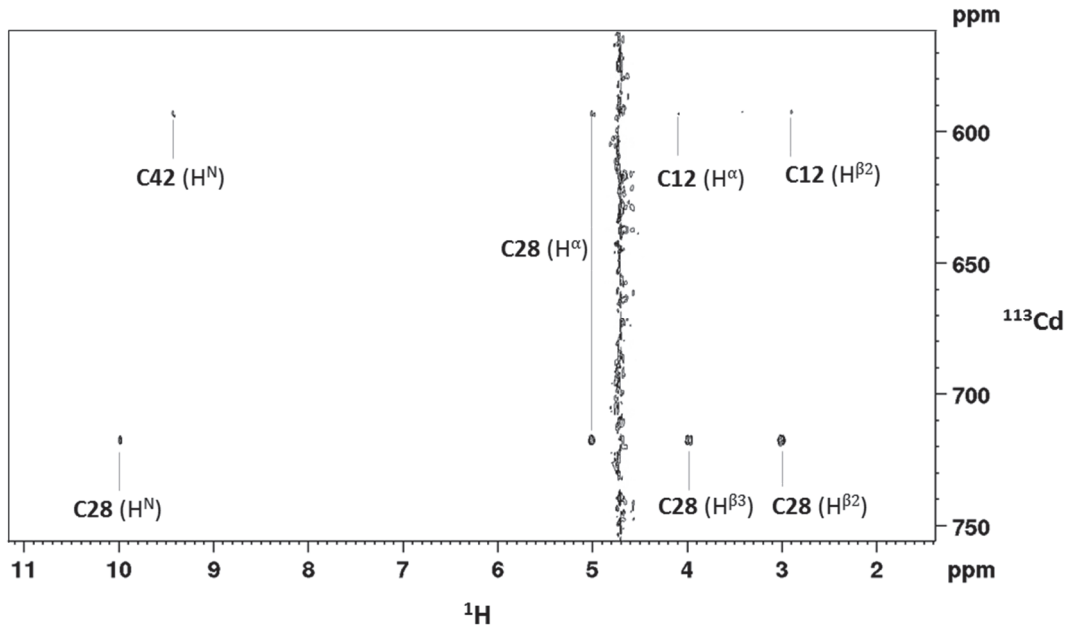

Figure S30. [^{113}Cd,^{1}H]-HSQC-TOCSY spectrum of ^{113}Cd$_4$A44H-sh_PflQ2 MT (T=320 K, J(^{113}Cd ^{1}H)= 40 Hz).

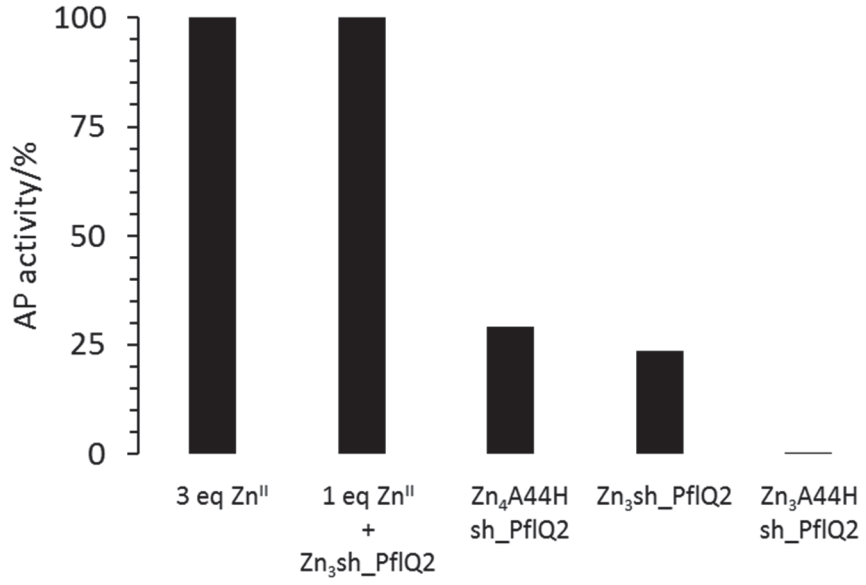

Figure S31. Zinc ion transfer to apo-alkaline phosphatase at t = 20 min.

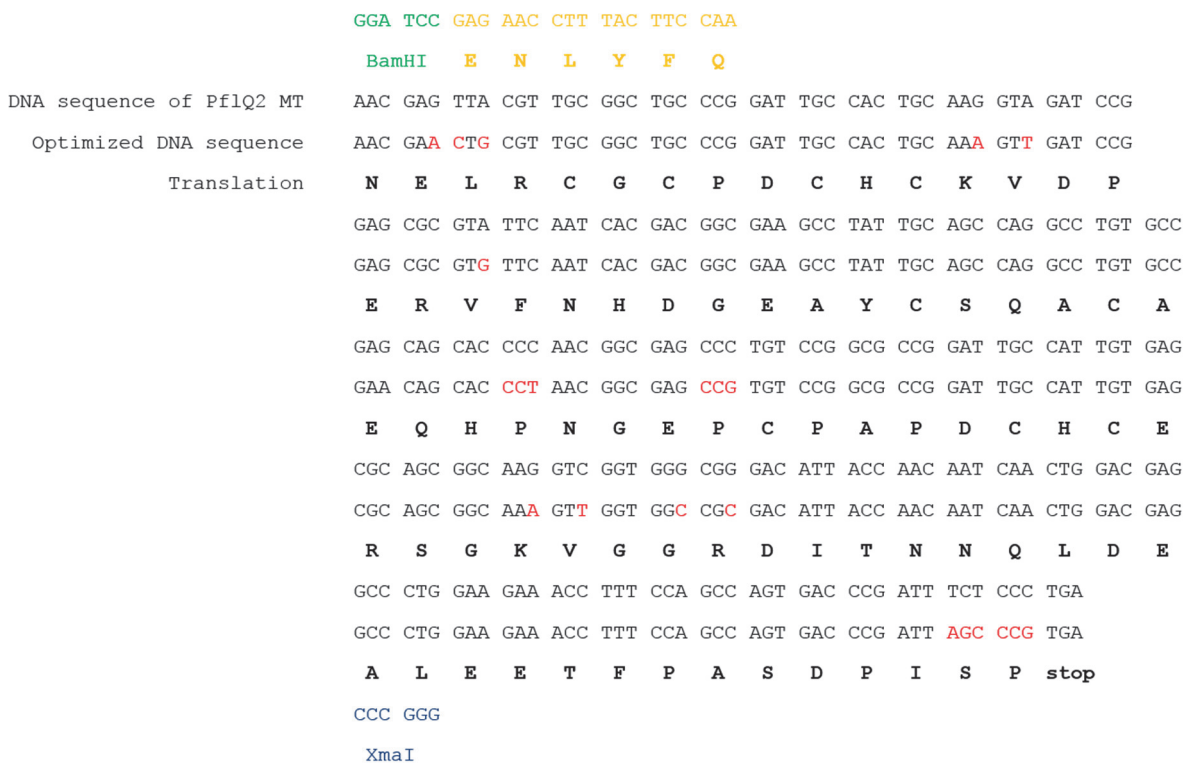

Figure S32. Codon usage for PflQ2 MT and its optimization for expression in *E. coli*.

Table S2. List of primers used for the construction of coding sequences.

J2EKT7_F1_fwd	GCG AAT TGG ATC CGA GAA CCT TTA CTT CCA AAA CGA ACT GCG TTG CGG CTG CCC GGA TTG
J2EKT7_R2_rev	GCC GTC GTG ATT GAA CAC GCG CTC CGG ATC AAC TTT GCA GTG GCA ATC CGG GCA GCC
J2EKT7_R3_rev	GTT AGG GTG CTG TTC GGC ACA GGC CTG GCT GCA ATA GGC TTC GCC GTC GTG ATT GAA C
J2EKT7_R4_rev	GCT GCG CTC ACA ATG GCA ATC CGG CGC CGG ACA CGG CTC GCC GTT AGG GTG CTG TTC G
J2EKT7_R5_rev	CGT CCA GTT GAT TGT TGG TAA TGT CGC GGC CAC CAA CTT TGC CGC TGC GTC ACA ATG G
J2EKT7_R6_rev	GTC ACT GGC TGG AAA GGT TTC TTC CAG GGC CTC GTC CAG TTG ATT GTT GG
J2EKT7_R7_rev	GTT CCC CCC GGG TCA CGG GCT AAT CGG GTC ACT GGC TGG AAA G
J2EKT7_sh_R4_rev	GCT CAC AAT GGC AAT CCG GCG CCG GAC ACG GCT CGC CGT TAG GGT GCT GTT CG
J2EKT7_sh_R5_rev	GTT CCC CCC GGG TCA GCT GCG CTC ACA ATG GCA ATC C
J2EKT7_sh_R4_A44H_rev	GCT CAC AAT GGC AAT CCG GAT GCG GAC ACG GCT CGC CGT TAG GGT GCT GTT CG
J2EKT7_sh_H11N_fwd	GCT GCC CGG ATT GCA ACT GCA AAG TTG ATC
J2EKT7_sh_H11N_rev	GAT CAA CTT TGC AGT TGC AAT CCG GGC AGC
J2EKT7_sh_H22N_fwd	CGC GTG TTC AAT AAC GAC GGC GAA GC
J2EKT7_sh_H22N_rev	GCT TCG CCG TCG TTA TTG AAC ACG CG
J2EKT7_sh_H48N_fwd	GCG CCG GAT TGC AAT TGT GAG CGC AG
J2EKT7_sh_H48N_rev	CTG CGC TCA CAA TTG CAA TCC GGC GC

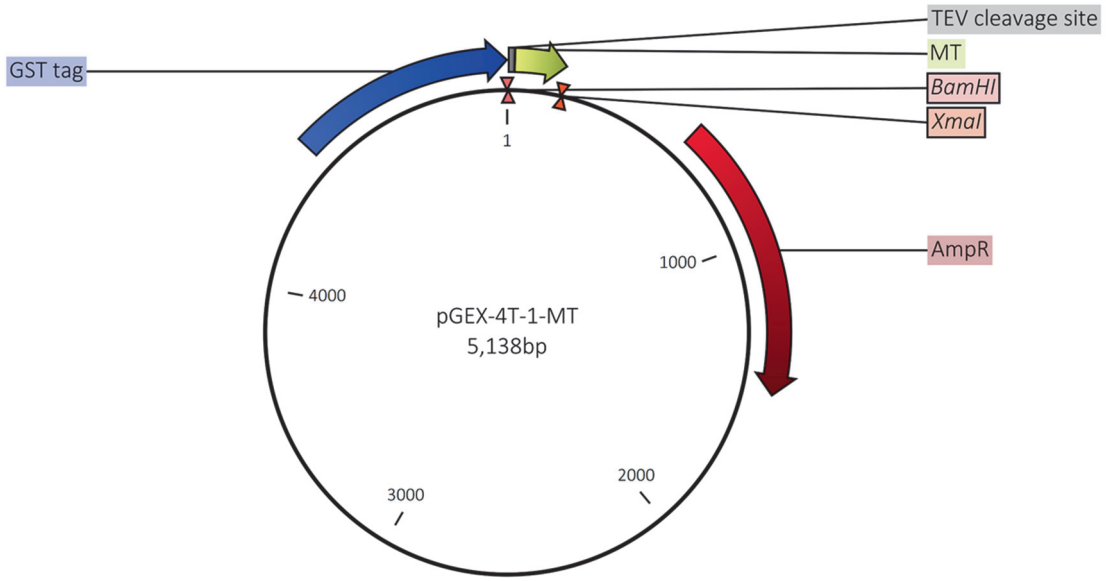

Figure S33. Plasmid map of pGEX-4T-1-MT expression vector.

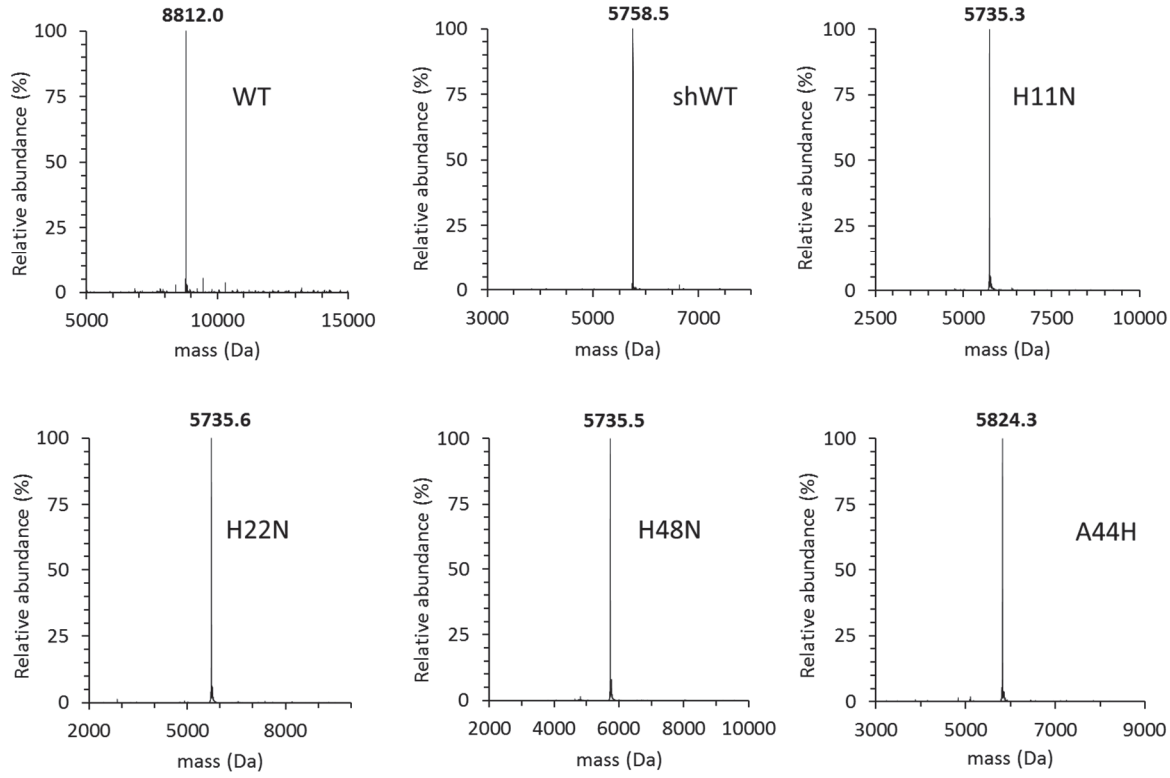

Figure S34. Deconvoluted ESI(+) mass spectra of apo-PflQ2 MT (WT) and its mutants. (WT (M_{calc} 8812.56 Da), shWT (M_{calc} 5759.28 Da), H11N, H22N, H48N (M_{calc} 5736.24 Da), A44H (M_{calc} 5825.34 Da)).

Table S3. Metal-ligand connectivities of Cd_4 and Cd_3 cluster in sh_PflQ2 MT. Connectivities obtained from [^{113}Cd, ^{1}H]-HSQC-TOCSY spectra are marked in black, those resulting from structure calculations in red. For the involvement of H36 in coordination see the main text.

Cd_4	Cd(A) (713.8 ppm)	Cd(B) (615.4 ppm)	Cd(C) (585.8 ppm)	Cd(D) (535 ppm)
	C5	**C10**	**C12**	**H36**
	C28	**C42**	**C28**	**C7**
	C32	**C47**	**C42**	**C32**
	C10	**C49**	**C10**	**C49**

Cd_3	Cd(A) (695 ppm)	Cd(B) (676 ppm)	Cd(D) (562 ppm)
	C5	**C10**	**H36**
	C28	**C42**	**C49**
	C12	**C47**	**C32**
	C32	**C49**	**C7**

Table S4. Details of the structure calculation and geometry of Cd_4PflQ2 MT, Cd_4sh_PflQ2 MT, Zn_3sh_PflQ2 MT and Cd_4A44H-sh_PflQ2 MT.

	Cd_4PflQ2 MT	
Input data for structure calculation		
NOE distance restraints		
total	1634	
short-range, i-j<2	1215	
medium range, 1>i-j>5	180	
long-range, i-j ≥5	239	
Structure statistics, 20 conformers		
CYANA target function value (Å^2)	0.78-0.82 (average: 0.80)	
XPLOR energies (kcal/mol)	TOTAL: -1316.25 ± 42.79	NOE: 23.36 ± 3.71
Ramachandran plot analysis (Procheck)	residues: all	residues: 3-50
Residues in favored regions (%)	78.5	85.3
Residues in additional allowed regions (%)	18.5	12.6
Residues in generously allowed regions (%)	1.5	2.2
Residues in disallowed regions (%)	1.5	0
Root mean square deviation to the average (Å)	residues: all	residues: 3-50
Average backbone RMSD to mean	10.708 ± 2.755	0.215 ± 0.033
Average heavy atom RMSD to mean	10.953 ± 2.771	0.660 ± 0.089
	Cd_4sh_PflQ2 MT	
Input data for structure calculation		
NOE distance restraints		
total	1143	
short-range, i-j<2	881	
medium range, 1>i-j>5	111	
long-range, i-j ≥5	151	
Structure statistics, 20 conformers		
CYANA target function value (Å^2)	0.48-0.52 (average: 0.52)	
XPLOR energies (kcal/mol)	TOTAL: -907.88 ± 53.45	NOE: 9.14 ± 2.13
Ramachandran plot analysis (Procheck)	residues: all	
Residues in favored regions (%)	75.6	
Residues in additional allowed regions (%)	19.5	
Residues in generously allowed regions (%)	4.9	
Residues in disallowed regions (%)	0.0	
Root mean square deviation to the average (Å)	residues: all	residues: 5-49
Average backbone RMSD to mean	0.781 ± 0.165	0.502 ± 0.134
Average heavy atom RMSD to mean	1.316 ± 0.194	0.965 ± 0.140

	Zn_3sh_PflQ2 MT	
Input data for structure calculation		
NOE distance restraints		
total	1354	
short-range, i-j<2	1030	
medium range, 1>i-j>5	110	
long-range, i-j ≥5	214	
Structure statistics, 20 conformers		
CYANA target function value (Å^2)	0.61-0.72 (average: 0.69)	
XPLOR energies (kcal/mol)	-1092.44 ± 60.34	NOE: 20.61 ± 6.57
Ramachandran plot analysis (Procheck)	residues: all	
Residues in favored regions (%)	85.4	
Residues in additional allowed regions (%)	14.6	
Residues in generously allowed regions (%)	0	
Residues in disallowed regions (%)	0	
Root mean square deviation to the average (Å)	residues: all	residues: 5-49
Average backbone RMSD to mean	0.957 ± 0.242	0.443 ± 0.066
Average heavy atom RMSD to mean	1.629 ± 0.248	0.973 ± 0.090

	Cd_4A44H-sh_PflQ2 MT	
Input data for structure calculation		
NOE distance restraints		
total	1485	
short-range, i-j<2	1099	
medium range, 1>i-j>5	120	
long-range, i-j ≥5	266	
Structure statistics, 20 conformers		
CYANA target function value (Å^2)	0.75.-0.96 (average: 0.86)	
XPLOR energies (kcal/mol)	-887.93 ± 64.79	NOE: 19.02 ± 6.21
Ramachandran plot analysis (Procheck)	residues: all	
Residues in favored regions (%)	80.5	
Residues in additional allowed regions (%)	17.1	
Residues in generously allowed regions (%)	0	
Residues in disallowed regions (%)	2.4	
Root mean square deviation to the average (Å)	residues: all	residues: 5-49
Average backbone RMSD to mean	0.909 ± 0.174	0.373 ± 0.068
Average heavy atom RMSD to mean	1.630 ± 0.184	0.933 ± 0.103

References

1 A. Bateman, M. J. Martin, C. O'Donovan, M. Magrane, E. Alpi, R. Antunes, B. Bely, M. Bingley, C. Bonilla, R. Britto, B. Bursteinas, H. Bye-A-Jee, A. Cowley, A. Da Silva, M. De Giorgi, T. Dogan, F. Fazzini, L. G. Castro, L. Figueira, P. Garmiri, G. Georghiou, D. Gonzalez, E. Hatton-Ellis, W. Z. Li, W. D. Liu, R. Lopez, J. Luo, Y. Lussi, A. MacDougall, A. Nightingale, B. Palka, K. Pichler, D. Poggioli, S. Pundir, L. Pureza, G. Y. Qi, S. Rosanoff, R. Saidi, T. Sawford, A. Shypitsyna, E. Speretta, E. Turner, N. Tyagi, V. Volynkin, T. Wardell, K. Warner, X. Watkins, R. Zaru, H. Zellner, I. Xenarios, L. Bougueleret, A. Bridge, S. Poux, N. Redaschi, L. Aimo, G. Argoud-Puy, A. Auchincloss, K. Axelsen, P. Bansal, D. Baratin, M. C. Blatter, B. Boeckmann, J. Bolleman, E. Boutet, L. Breuza, C. Casal-Casas, E. de Castro, E. Coudert, B. Cuche, M. Doche, D. Dornevil, S. Duvaud, A. Estreicher, L. Famiglietti, M. Feuermann, E. Gasteiger, S. Gehant, V. Gerritsen, A. Gos, N. Gruaz-Gumowski, U. Hinz, C. Hulo, F. Jungo, G. Keller, V. Lara, P. Lemercier, D. Lieberherr, T. Lombardot, X. Martin, P. Masson, A. Morgat, T. Neto, N. Nouspikel, S. Paesano, I. Pedruzzi, S. Pilbout, M. Pozzato, M. Pruess, C. Rivoire, B. Roechert, M. Schneider, C. Sigrist, K. Sonesson, S. Staehli, A. Stutz, S. Sundaram, M. Tognolli, L. Verbregue, A. L. Veuthey, C. H. Wu, C. N. Arighi, L. Arminski, C. M. Chen, Y. X. Chen, J. S. Garavelli, H. Z. Huang, K. Laiho, P. McGarvey, D. A. Natale, K. Ross, C. R. Vinayaka, Q. H. Wang, Y. Q. Wang, L. S. Yeh, J. Zhang and U. Consortium, UniProt: the universal protein knowledgebase, *Nucleic Acids Res.*, 2017, **45**, D158-D169.

2 M. E. Oates, P. Romero, T. Ishida, M. Ghalwash, M. J. Mizianty, B. Xue, Z. Dosztanyi, V. N. Uversky, Z. Obradovic, L. Kurgan, A. K. Dunker and J. Gough, D^2P^2: Database of disordered protein predictions, *Nucleic Acids Res.*, 2013, **41**, D508-D516.

3 C. A. Blindauer, Bacterial metallothioneins, *Met. Ions Life Sci.*, 2009, **5**, 51-81.