De Novo Design of an Allosteric Metalloprotein Assembly with Strained Disulfide Bonds

Lewis A. Churchfield, Annette Medina-Morales, Jeffrey D. Brodin, Alfredo Perez, and F. Akif Tezcan*

Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093-0356.

This file contains Supporting Tables S1-4, and Supporting Figures S1-S9.

Table S1. Thermodynamic parameters for Zn^{2+} binding to disulfide-linked $R1_4$ tetramers. Binding affinities were determined by Zn^{2+} titration in the presence of Fura-2. Errors denote fitting errors of a single measurement. Binding enthalpies were measured by ITC for the titration of Zn^{2+} with disulfide-linked $R1_4$ variants. Reported enthalpies are not corrected for buffer effects, and reported errors are the standard deviation among 3-4 replicate measurements. All measurements were carried out at 22 °C and in the presence of 20 mM MOPS, pH 7 + 150 mM NaCl. [a]Data taken from reference 1. [b]Data taken from reference 2.

	$^{C96}R1_4$	$^{C81/C96}R1_4$	$^{C38/C81/C96}R1_4$
Binding affinities			
K_{d1} (nM)	1.3 ± 0.3 [a]	2.6 ± 0.3 [b]	8.1 ± 0.4
K_{d2} (nM)	0.53 ± 0.07 [a]	25 ± 4 [b]	400 ± 100
K_{d3} (nM)	33 ± 8 [a]	-	-
K_{d4} (nM)	58 ± 8 [a]	-	-
Binding enthalpies			
$\Delta H_{1,ITC}$ (kJ/mol)	18 ± 3	14 ± 4	4 ± 2
$\Delta H_{2,ITC}$ (kJ/mol)	14 ± 10	33 ± 6	5 ± 3
$\Delta H_{3,ITC}$ (kJ/mol)	33 ± 15	11 ± 10	2 ± 3
$\Delta H_{4,ITC}$ (kJ/mol)	−3 ± 4	1 ± 2	−1 ± 2

Table S2. X-ray crystallography data collection and refinement statistics. Values in parentheses denote statistics for the highest resolution shell.

	$^{C38/C81/C96}$**R1**$_4$	**Zn-**$^{C38/C81/C96}$**R1**$_4$
Data collection		
Space group	$P\,2_12_12_1$	$P2_1$
Cell dimensions		
a, *b*, *c* (Å)	62.59, 77.53, 88.06	48.06, 62.53, 72.44
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 99.01, 99.00
Resolution (Å)	38.77-2.40	37.81-2.10
R_{sym} or R_{merge} (%)	6.7 (23.0)	4.9 (46.3)
I/σ	14.9 (5.7)	15.0 (3.0)
$CC_{1/2}$ (%)	99.9 (96.8)	99.9(79.6)
Completeness (%)	99.8 (99.6)	99.5 (99.3)
Redundancy	6.9	3.7
Refinement		
Resolution (Å)	2.40	2.10
No. unique reflections	17325	24815
R_{work}/R_{free} (%)	22.5/26.5	21.4/25.0
No. atoms		
Protein	3224	3264
Ligand/ion	172	195
Water	64	73
***B*-factors**		
Protein	56.54	47.19
Ligand/ion	54.31	43.45
Water	47.83	39.94
R.m.s. deviations		
Bond lengths (Å)	0.008	0.018
Bond angles (°)	1.30	1.57

Table S3. Calculated strain energies of disulfide bonds in the available $^{C83/C81/C96}R1_4$ structures. Reported energies are averaged over disulfide pairs within each structure, with the exception of the single C38-C38 disulfide in apo-$^{C83/C81/C96}R1_4$. Strain energies are calculated previously described by Katz and Kossiakoff.[3]

Disulfide site	Average strain energy (kJ/mol)	
	$^{C38/C81/C96}R1_4$	Zn-$^{C38/C81/C96}R1_4$
C38-C38	25	19
C81-C81	23	13
C96-C96	17	6

Table S4. Relative accessible surface areas of cysteine residues in the available $^{C38/C81/C96}R1_4$ structures. Reported values are averaged over disulfide pairs within each structure, with the exception of the single C38-C38 disulfide in apo-$^{C83/C81/C96}R1_4$. Values were calculated using the ASAView server (http://www.abren.net/asaview/).[4]

Disulfide site	Relative accessible surface areas	
	$^{C38/C81/C96}R1_4$	Zn-$^{C38/C81/C96}R1_4$
C38-C38	0.08	0.10
C81-C81	0.21	0.29
C96-C96	0.14	0.21

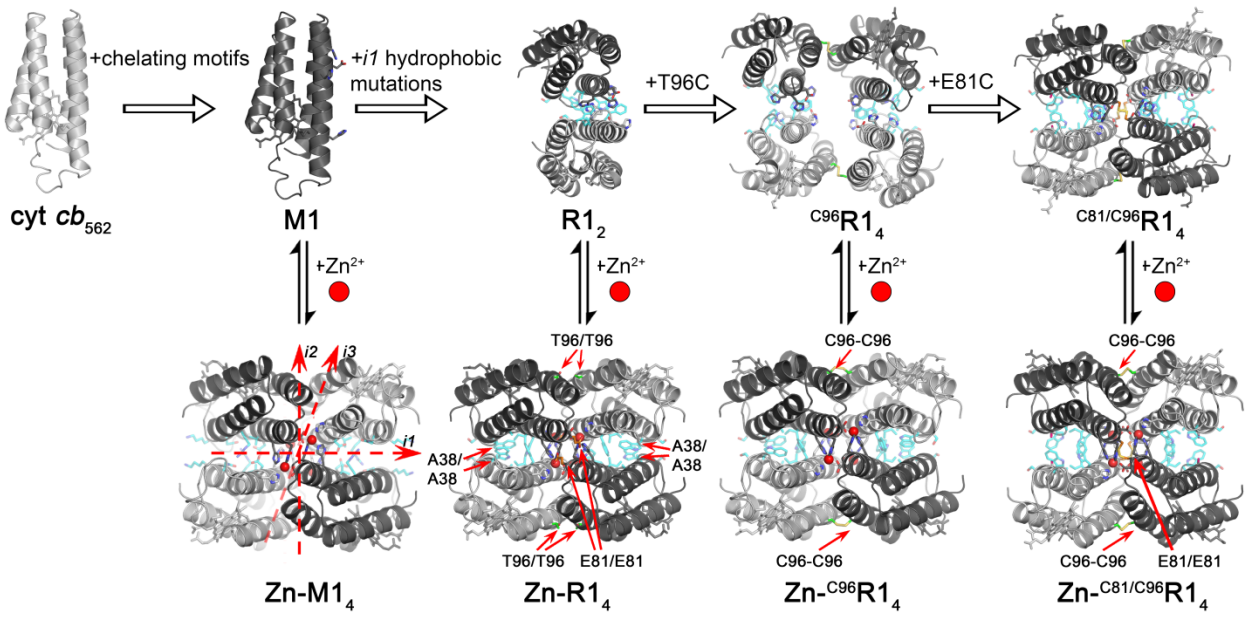

Figure S1. Successive engineering of cyt cb_{562} to form disulfide-linked oligomeric architectures. Using Metal Templated Interface Redesign (MeTIR), metal chelating motifs (shown as sticks) were installed on the surface of monomeric cyt cb_{562} (PDB ID: 2BC5). The resulting variant, M1 (or MBPC1, PDB ID: 2QLA), forms a Zn-dependent tetramer, Zn-M1$_4$. This tetrameric scaffold contains three C_2 symmetric interfaces, *i1*, *i2*, and *i3*. The scaffold was further engineered to include favorable protein-protein contacts at interface *i1* through the installation of hydrophobic patches (cyan)**.** The second-generation variant, R1, self assembles into a dimer (R1$_2$, PDB ID: 3HNK) and forms a Zn-induced tetramer (Zn-R1$_4$, PDB ID: 3HNI) that is topologically identical to the parent variant. Within the tetramer are symmetry-related residues, A38/A38, E81/E81, and T96/T96, that are in close proximity. Previous work has shown that Zn-mediated R1 assembly can template C96-C96 (green sticks) disulfide bond formation to afford a tetrameric species that is maintained in the absence of metal (^{C96}R1$_4$, PDB IDs: 3IQ5 and 3IQ6). Additionally, in the case of $^{C81/C96}$R1$_4$, Zn-mediated R1 assembly can template the formation of C96-C96 and C81-C81 (orange sticks) disulfide bonds simultaneously (PDB IDs: 4JE9 and 4JEA).

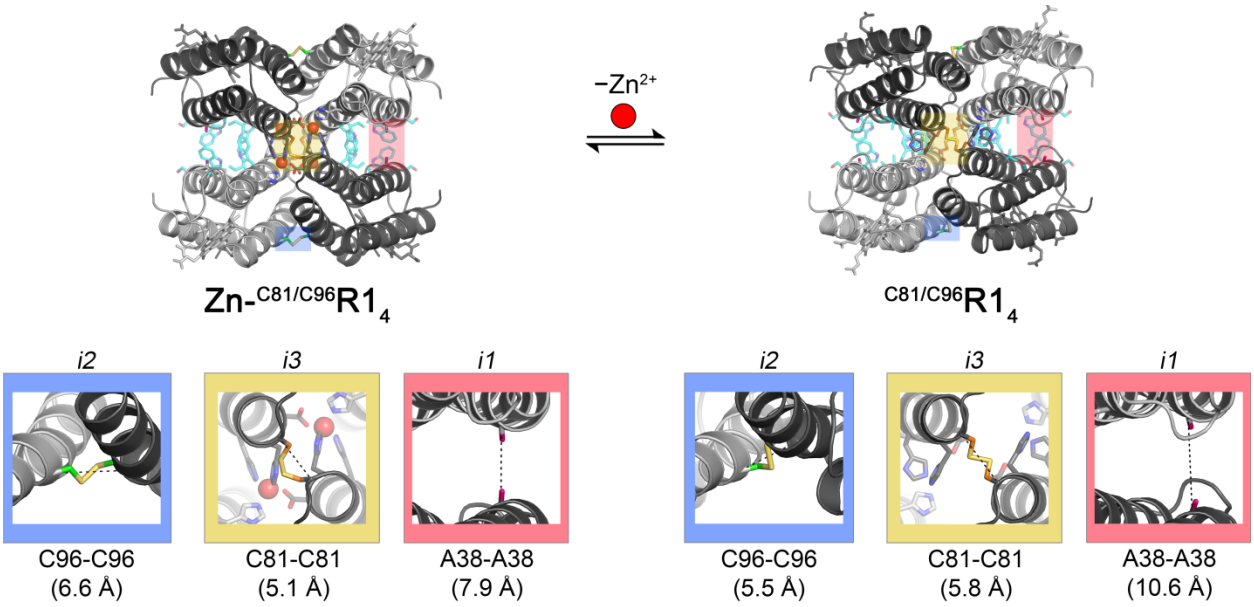

Figure S2. Structural rearrangements upon Zn^{2+} removal from Zn-$^{C81/C96}R1_4$. The disulfide crosslinks across interfaces *i2* and *i3* (C96-C96 and C81-C81, respectively), undergo conformational changes in a hinge-like fashion. The accompanying structural rearrangements increase the separation of the symmetry-related A38-A38 residues from 7.9 Å to 10.6 Å. Dashed lines and values in parentheses denote the average α-carbon separations of the indicated residues in the Zn- and apo-$^{C38/C81/C96}R1_4$ structures.

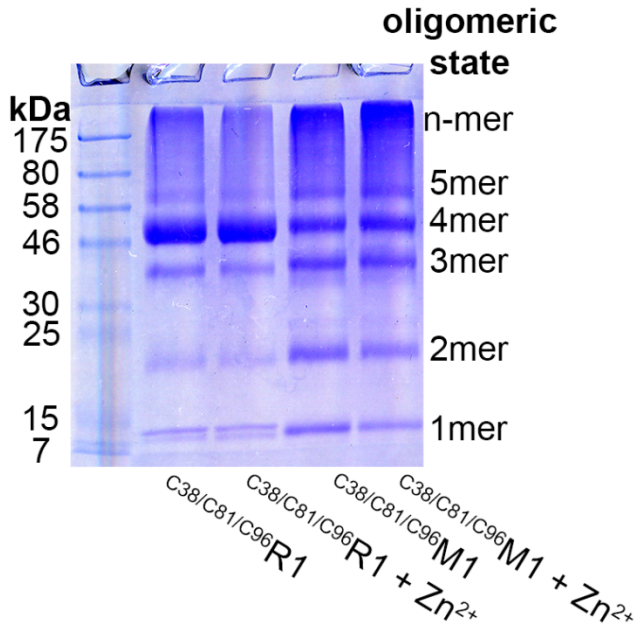

Figure S3. SDS-PAGE gel of the products of $^{C38/C81/C96}$R1 and $^{C38/C81/C96}$M1 self-assembly reactions in the presence or absence of Zn^{2+}. The $^{C38/C81/C96}$M1 variant lacks the hydrophobic mutations engineered to stabilize the *i1* interfaces in $^{C38/C81/C96}$R1. The gel was run in the absence of any reductants to keep disulfide bonds intact. Self-assembly conditions are described in Section 2.2 of the Supplementary Methods.

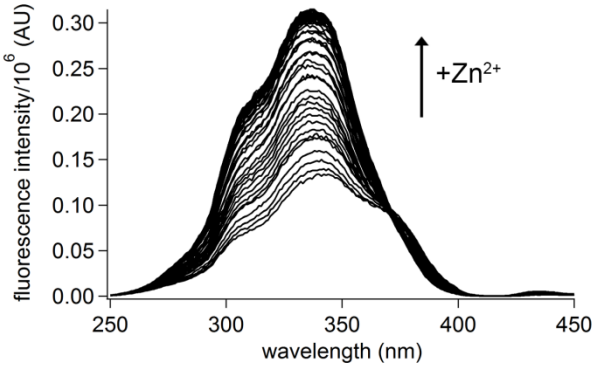

Figure S4. Competitive Zn^{2+} binding titration of $^{C38/C81/C96}R1_4$ in the presence of Fura-2. Zn competition was monitored by excitation scans at a fixed emission wavelength of 510 nm. The corresponding changes in the intensity at 335 nm (shown above) were used to generate the Zn binding isotherms shown in Figure 3.

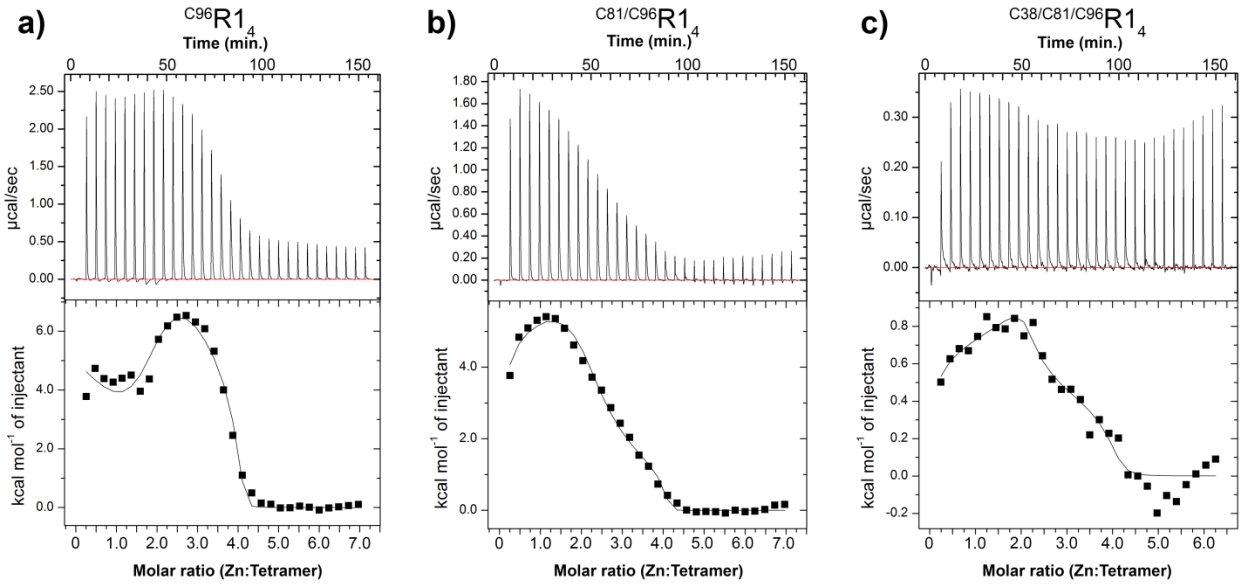

Figure S5. Representative baseline-corrected and integrated ITC thermograms for disulfide-linked $R1_4$ variants titrated with $ZnCl_2$. ITC titrations and data fitting were carried out as described in Section 2.6 of the Supplementary Methods.

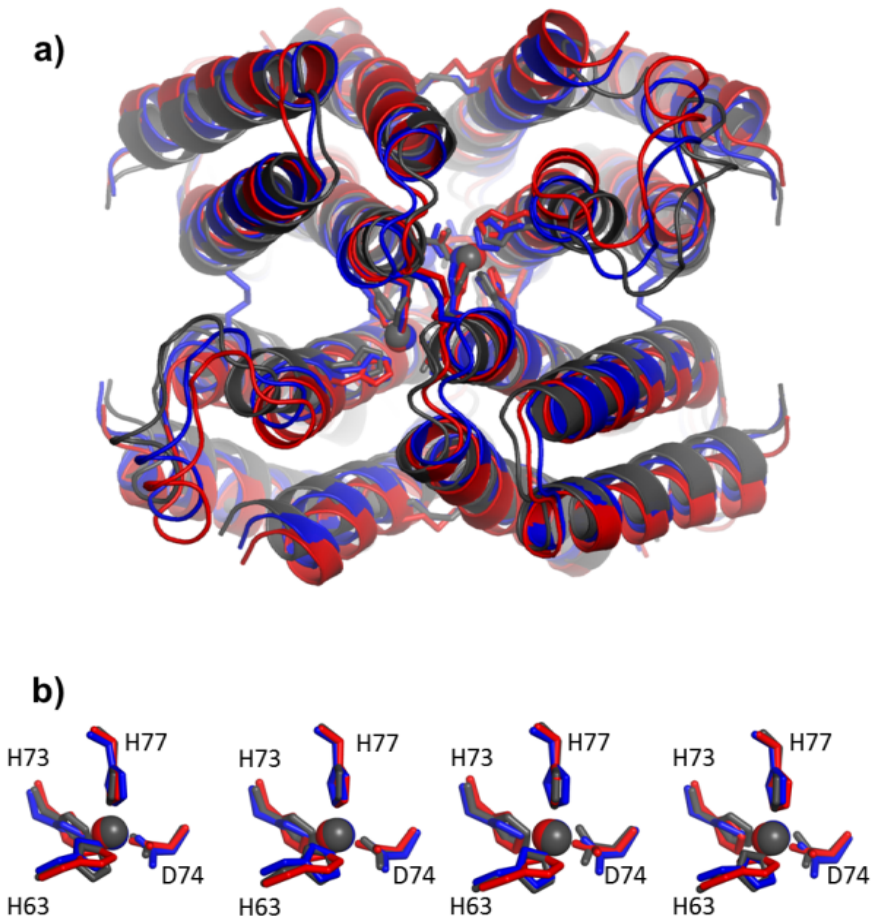

Figure S6. (**a**) Structural overlay of Zn-^{C96}R1$_4$ (black), Zn-$^{C81/C96}$R1$_4$ (red), and Zn-$^{C38/C81/C96}$R1$_4$ (blue) aligned over all α-carbon atoms. (**b**) Individual Zn^{2+} sites within each structure aligned to the ligand atoms only.

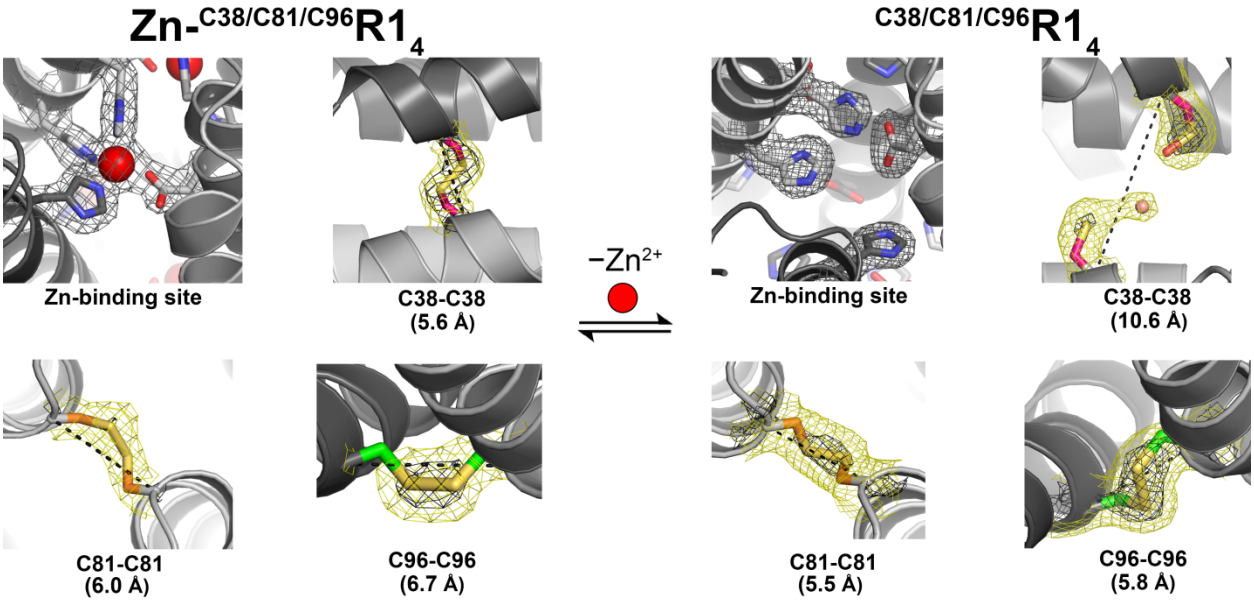

Figure S7. Close-up views of the Zn^{2+} coordination sites and disulfide bonds in Zn-$^{C38/C81/C96}R1_4$ and $^{C38/C81/C96}R1_4$ structures. $2F_o$-F_c electron density maps are contoured at 1 σ (yellow) and 2.5 σ (black). Dashed lines and values in parentheses denote the α-carbon separation between the indicated residue pairs.

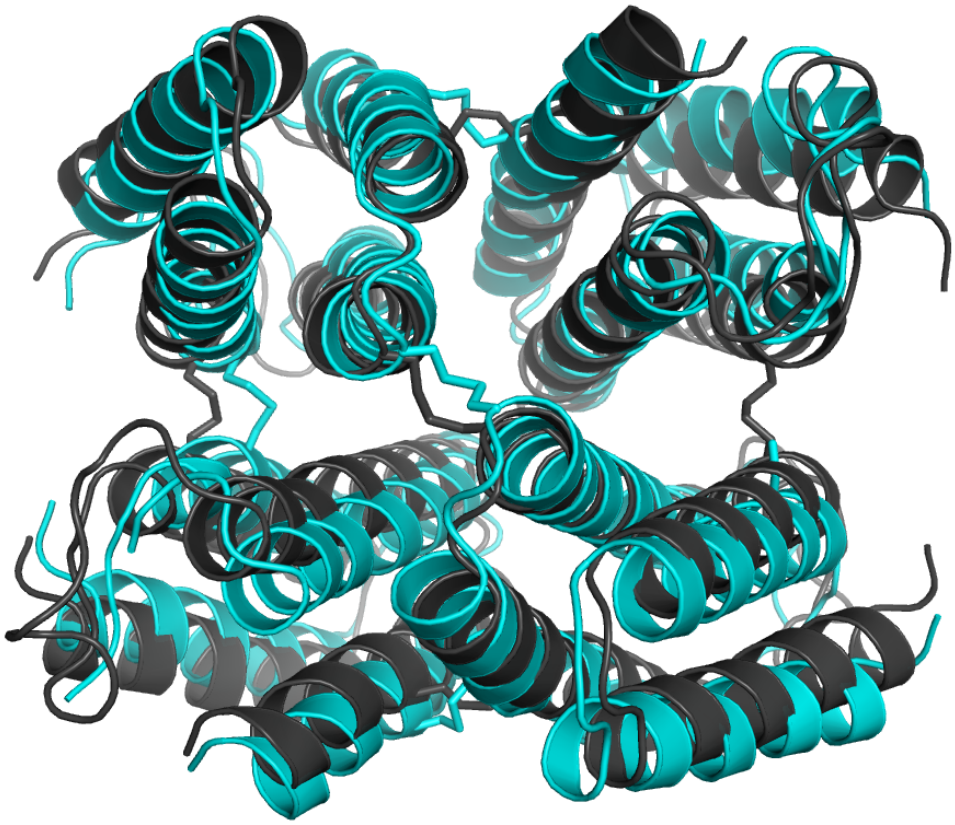

Figure S8. Structural overlay of Zn-$^{C38/C81/C96}$R1$_4$ (black) and $^{C38/C81/C96}$R1$_4$ (cyan) assemblies aligned over all α-carbon atoms (RMSD = 2.55 Å). Disulfide bonds and the broken C38-C38 disulfide are shown as sticks.

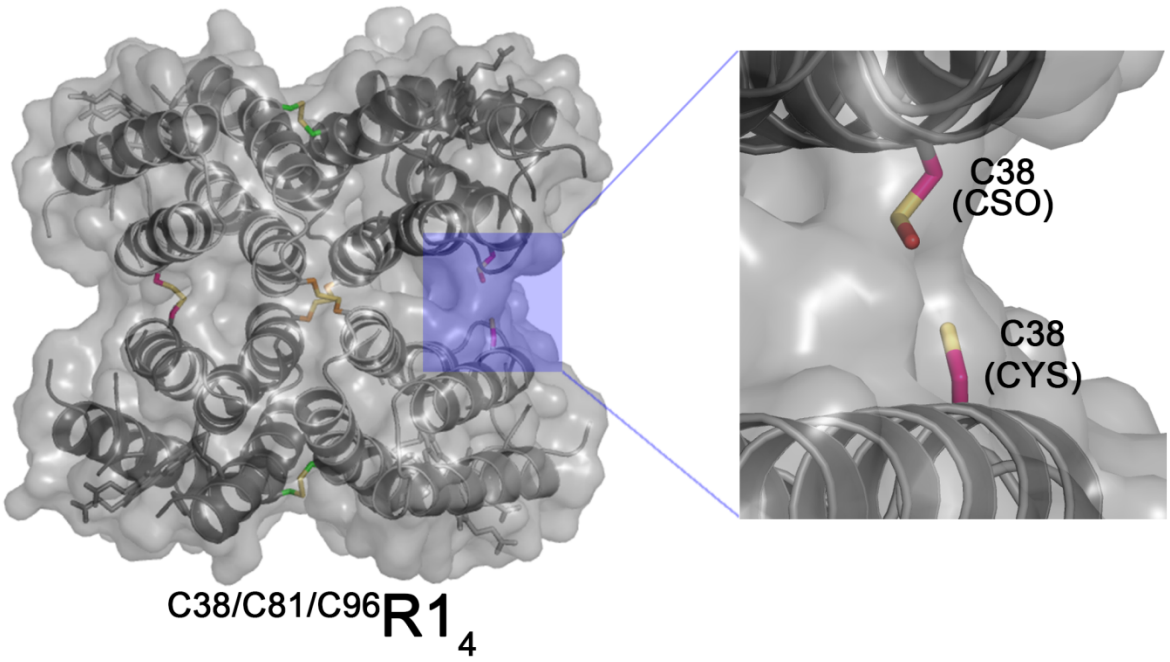

Figure S9. Surface representation of the $^{C38/C81/C96}R1_4$ crystal structure contoured at 1.4 Å. The close-up view shows the broken C38-C38 disulfide bond, which is located within an interfacial cleft of the tetramer. Disulfide bonds and the broken C38-C38 disulfide are shown as sticks.

Supplementary Methods

2.1 Protein sample preparation. All constructs used were housed in a pet20b vector (Novagen), and coding sequences included the N-terminal periplasmic localization sequence of *R. palustris* cytochrome c_{556}.[5] The A38C and L38C mutations were introduced into vectors (Novagen) harboring the genes encoding $^{C81/C96}$R1 and $^{C81/C96}$M1, respectively, by site-directed mutagenesis using oligonucleotide primers obtained from Integrated DNA Technologies. The primer pair 5'-cgctccacgcatcgcacgctgcggccgc-3' and 5'-ggcggccgcagcgtgcgatgcgtggagcg-3' was used with the pET20b/$^{C81/C96}$R1 template to yield the gene encoding $^{C38/C81/C96}$R1, and the primer pair 5'-gcgcgccgcagcgtgcgatgcgtggagcg-3' and 5'-cgctccacgcatcgcacgctgcggcgcgc-3' was used with the pET20b/$^{C81/C96}$M1 template to yield the gene encoding $^{C38/C81/C96}$M1. DNA amplification was carried out using PfuTurbo DNA Polymerase (Agilent Technologies). The amplified DNA was transformed into XL-1 Blue competent *E. coli* cells (Agilent Technologies), and the cells were grown at 37 °C on lysogeny broth (LB) agar plates containing 60 μg/mL ampicillin. Plasmids were purified with the QIAprep Spin Miniprep Kit (QIAGEN), verified by sequencing (Retrogen), and then transformed into BL21(DE3) *E. coli* cells (New England Biolabs) housing the cytochrome c maturation (ccm) plasmid cassette.[6] The cells were grown at 37 °C on LB agar plates containing 34 μg/mL chloramphenicol and 60 μg/mL ampicillin. Single colonies were used to inoculate 5 mL of liquid LB medium containing chloramphenicol and ampicillin. Cultures were shaken at 37 °C until visibly turbid (typically about 8 h, or to an OD_{600}>0.6) at which time 50 μL of inoculum was transferred to 1 L of LB medium, typically to 15 cultures in parallel, and shaken for 16-20 h at 37 °C with protein expression occurring by auto-induction. Cells were harvested by centrifugation (4,000 rpm, 4 °C, 10 min), resuspended in 100 mL of 5 mM sodium acetate (pH 5), frozen, thawed, and sonicated for 12 min in pulses of 30 s on and 60 s off in the presence of lysozyme and excess DTT while on ice. The lysate was titrated with sodium hydroxide to pH >10 and acetic acid to pH 5, and cleared by centrifugation (10,000 rpm, 4 °C, 10 min). The red cleared lysate was decanted and diluted to 2 L in 5 mM sodium acetate (pH 5) containing 2 mM DTT, and was manually applied to a CM Sepharose column (GE Healthcare). The sample was washed with 5 mM sodium acetate containing 2 mM DTT (pH 5), and eluted in a 0-1 M gradient of NaCl. The clear red eluate was concentrated and buffer-exchanged into 10 mM sodium phosphate (pH 8) containing 2 mM DTT using a Diaflow concentrator (Amicon) fitted with a 3-kDa cutoff membrane. The sample was loaded onto a DuoFlow fast protein liquid chromatography station fitted with a Macroprep High Q-cartridge column (BioRad) with 10 mM sodium phosphate (pH 8) running buffer containing 2 mM DTT, and eluted using a 0-0.5M NaCl gradient. Fractions with Reinheitzahl values (A_{421}/A_{280}) above 3 were combined and concentrated, and excess EDTA and DTT (>10 molar equivalents) was added to the protein samples to remove any bound metal ions and prevent disulfide formation. These stock solutions were then flash frozen in liquid nitrogen for storage at −80 °C.

2.2 Tetramer self-assembly and purification. Concentrated stock solutions of $^{C38/C81/C96}$R1 (or $^{C38/C81/C96}$M1) were thawed, reduced by the addition of excess DTT, transferred to glass vials sealed with rubber septa, cycled under vacuum and argon atmosphere, and transferred to a glove box (Coy Lab) under an anaerobic (<10 ppm O_2) atmosphere of argon with 10% hydrogen. Here, the samples were exchanged into an assembly buffer solution (50 mM Tris, pH 7, and 150 mM NaCl) using Econo-Pac 10DG pre-packed columns (BioRad). Protein concentration was determined spectrophotometrically ($\varepsilon_{421,\,red}$ = 162,000 M^{-1} cm^{-1}; $\varepsilon_{421,\,ox}$ = 148,000 M^{-1} cm^{-1}),[5] and samples were diluted to a final concentration of 50 μM in assembly buffer supplemented with 75 μM $ZnCl_2$. Samples were incubated overnight under ambient atmosphere at 37 °C with shaking. Self-assembly reactions were analyzed by SDS-PAGE in the

absence of added reductant to keep disulfide bonds intact. The crude self-assembly reaction mixture was applied to a preparative scale Superdex 75 gel filtration column (GE Healthcare) equilibrated in assembly buffer. Purities of the tetramer-containing fractions were assessed by non-reducing SDS-PAGE. The purest fractions were combined, concentrated, and treated with >10-fold excess EDTA to remove bound metal, and applied to an Econo-Pac 10DG prepacked column equilibrated in 20 mM MOPS (pH 7) and 150 mM NaCl. The resulting stock solutions of Zn-free tetramer were flash frozen in liquid nitrogen, and stored at −80 °C.

2.3 Sedimentation velocity analytical ultracentrifugation. Solutions of 1.25 μM tetramer ($^{C38/C81/C96}R1_4$) in 20 mM MOPS (pH 7) and 150 mM NaCl were treated with either 5 μM $ZnCl_2$ or 1 mM EDTA to prepare metallated or apo samples. Sedimentation velocity measurements were made on a XL-1 Analytical Ultracentrifuge (Beckman-Coulter) equipped with an An-60 Ti rotor at 41,000 rpm for 400 scans at 25 °C, and monitored at 415 nm. The endpoint of sedimentation was determined using the match function in HeteroAnalysis (http://biotech.uconn.edu/auf/). Scans were processed in Sedfit[7] using buffer viscosity (0.01002 poise), density (1.007 g/mL) and partial specific volume (0.7316 ml/g) parameters calculated by SEDNTERP (http://sednterp.unh.edu/). The final c(S) distributions are reported at a confidence level of 0.95.

2.4 Competitive Zn^{2+} titration assay. The concentration of a light-protected stock solution of Fura-2 (Invitrogen) was determined spectrophotometrically (ε_{362} = 27,000 M^{-1} cm^{-1}).[8] Samples of $^{C38/C81/C96}R1_4$ and Fura-2 were prepared at a final concentration of 7.5 μM in a buffered solution of 20 mM MOPS (pH 7) and 150 mM NaCl treated with Chelex 100 resin (BioRad). The sample was titrated with a $ZnCl_2$ solution while thermostatted at 22 °C, and fluorescence measurements were made after 5-min equilibration periods. Fura-2 fluorescence emission at 510 nm was monitored to obtain an excitation scan over 250-450 nm on a Horiba Fluorolog 2 fluorimeter (**Figure S4**). Binding isotherms were generated from the changes in emission intensity as a function of Zn^{2+} concentration, and were fit using Dynafit.[9] Measured dissociation constants are reported in **Table S1**, and total ΔG_{Zn} in **Table 1**.

2.5 X-ray crystallography and structural analysis. For crystallization of apo-$^{C38/C81/C96}R1_4$, a stock of 700 μM tetramer was prepared in 20 mM MOPS, pH 7, with 150 mM NaCl. For crystallization of Zn-$^{C38/C81/C96}R1_4$, a 375 μM stock of apo-protein in assembly buffer (20 mM Tris, pH 7, with 150 mM NaCl) was prepared and preincubated with 2250 μM $ZnCl_2$. Screens were carried out by sitting-drop vapor diffusion at room temperature with 500 μL reservoirs and droplets consisting of 1.5 μL of protein and 1 μL of precipitant solution. Crystals of metal-free $^{C38/C81/C96}R1_4$ were obtained with a precipitant solution consisting of 0.1 M Bis-Tris (pH 6.5) and 45% 2-methyl-2,4-pentanediol, and crystals of Zn-$^{C38/C81/C96}R1_4$ were obtained with a precipitant consisting of 0.2 M $MgCl_2$, 0.1 M Bis-Tris (pH 6.5), and 35% PEG 400. Crystals were harvested, cryoprotected in perfluoropolyether cryo oil (Hampton Research), and stored in liquid nitrogen. X-ray diffraction data for apo-$^{C38/C81/C96}R1_4$ and Zn-$^{C38/C81/C96}R1_4$ were collected at 100 K on beamlines BL9-2 and BL14-1, respectively, at the Stanford Synchrotron Radiation Lightsource. Diffraction data were integrated using Web-Ice.[10] Integrated datasets were scaled using SCALA[11] within the CCP4 suite. Molecular replacement was carried out using Phaser[12] with a $^{C96}R1$ monomer (PDB ID: 3IQ6) as the search model. Data refinement and model building were carried out using REFMAC[13] and COOT[14]. Four-fold non-crystallographic symmetry (NCS) restraints were applied in the initial stages of refinement and gradually relaxed. Ramachandran plots were

calculated using PROCHECK.[15] All structure figures and structural alignments were generated using PyMOL (www.pymol.org).

2.6 Isothermal titration calorimetry. Experiments were carried out on a VP-ITC instrument (MicroCal) with a cell volume of 1.4125 mL. Samples of $^{C96}R1_4$, $^{C81/C96}R1_4$ and $^{C38/C81/C96}R1_4$ were buffer-exchanged into a solution of 20 mM MOPS (pH 7) and 150 mM NaCl previously treated with Chelex-100 Resin (BioRad). Protein samples were then dialyzed twice at 4 °C for >8 h against the indicated buffer, with the second dialysis buffer retained for diluting protein samples and preparing $ZnCl_2$ titrant solutions. The calorimeter cell contained 25-50 μM of apo-tetramer and a water reference. The cell was equilibrated at 22 °C for 180 s before an initial injection of 1 μL of 1.4 mM $ZnCl_2$ titrant, followed by a schedule of 30 injections of 5.5 μL $ZnCl_2$ delivered at 5 min intervals. Data analysis was carried out using the ORIGIN data analysis plugin from MicroCal. The initial injection was omitted from data analysis. Integrated data were background-subtracted using the final injections of the thermogram to account for metal dilution effects, and were fit to a four-site sequential binding model. ITC thermograms were fit using dissociation constants determined by fluorescence competition assays, and enthalpy values were allowed to vary. The dissociation constants could not be directly determined by ITC owing to the tight binding constants (requiring ~40 μM protein to measure the weakest binding events, and even lower concentrations for tighter binding) and low enthalpies of binding (requiring ~50 μM binding site to achieve 1 μcal/injection). Therefore, the reported experimental conditions gave a value of $c > 1000$ (where c is the concentration of binding sites divided by the dissociation constant). This limitation has also been reported in other metalloprotein systems with high metal-binding binding affinities, where an ancillary experiment (e.g., equilibrium dialysis) can be used instead to determine the dissociation constants.[16] As previously noted,[16] the use of two independent experiments has the added benefit of avoiding correlated errors in determining ΔG and ΔH from measuring both parameters by ITC. Finally, we note that the complexity of Zn^{2+} binding to the $R1_4$ scaffolds (four ΔH_{ITC}'s and four K_d's, giving eight free parameters in total) necessitate a complex model that may not robustly measure these binding parameters on a per-site basis. Therefore, we employed an analysis of the global binding parameters supported by additional experiments. Representative raw and integrated thermograms, as well as their best-fit curves, are reported in **Figure S5**, and the measured ΔH_{ITC} values in **Table S1**.

References

(1) Brodin, J. D.; Medina-Morales, A.; Ni, T.; Salgado, E. N.; Ambroggio, X. I.; Tezcan, F. A. *J. Am. Chem. Soc.* **2010**, *132*, 8610.
(2) Medina-Morales, A.; Perez, A.; Brodin, J. D.; Tezcan, F. A. *J. Am. Chem. Soc.* **2013**, *135*, 12013.
(3) Katz, B. A.; Kossiakoff, A. *J. Biol. Chem.* **1986**, *261*, 15480.
(4) Ahmad, S.; Gromiha, M.; Fawareh, H.; Sarai, A. *BMC Bioinf.* **2004**, *5*, 51.
(5) Faraone-Mennella, J.; Tezcan, F. A.; Gray, H. B.; Winkler, J. R. *Biochemistry* **2006**, *45*, 10504.
(6) Arslan, E.; Schulz, H.; Zufferey, R.; Kunzler, P.; Thony-Meyer, L. **1998**, *251*, 744.
(7) Schuck, P. *Biophys J* **2000**, *78*, 1606.
(8) Grynkiewicz, G.; Poenie, M.; Tsien, R. Y. *J. Biol. Chem.* **1985**, *260*, 3440.
(9) Kuzmic, P. *Anal Biochem* **1996**, *237*, 260.
(10) Gonzalez, A.; Moorhead, P.; McPhillips, S. E.; Song, J.; Sharp, K.; Taylor, J. R.; Adams, P. D.; Sauter, N. K.; Soltis, S. M. *J Appl Crystallogr* **2008**, *41*, 176.

(11) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. **2011**, *67*, 235.
(12) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658.
(13) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr. D Biol. Crystallogr.* **1997**, *53*, 240.
(14) Emsley, P.; Cowtan, K. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126.
(15) Laskowski, R. A.; Macarthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283.
(16) DiTusa, C. A.; Christensen, T.; McCall, K. A.; Fierke, C. A.; Toone, E. J. *Biochemistry* **2001**, *40*, 5338.