# Supporting Information

## Salgado et al. 10.1073/pnas.0906852107

### SI Text

**Experimental Methods** All cyt $cb_{562}$ constructs were generated, expressed, and purified according to published procedures (1–3). Similarly, sedimentation velocity and equilibrium measurements/analyses, as well as protein unfolding titrations were carried out as previously described (1–3). More experimental detail can be found in corresponding *SI Text* figure captions.

X-band EPR spectra were obtained at 125 K on a Bruker ELEXSYS E500 spectrometer equipped with a Bruker ER4131VT variable-temperature unit. The EPR samples consisted of 150 μM RIDC-1 in 20 mM MOPS buffer (pH 7), with or without 100 μM copper (II) chloride dihydrate The spectrum of the copper-free sample was subtracted from that of the copper-containing sample to obtain the final spectrum. Spectra were recorded using the following spectrometer conditions: Microwave frequency, 9.389688 GHz; power, 0.63 mW; modulation amplitude, 0.10 mT.

All crystals were obtained by sitting drop vapor diffusion at room temperature. The crystallization conditions for the four different crystal forms described in this study are as follows: $Zn_4$ : RIDC–$1_4$ – 1 $\mu L$ of precipitant solution (100 mM HEPES, pH 7.0, 30% Jeffamine ED-2001 (pH 7.0), 1.5 mM $ZnCl_2$) and 2 μL protein (1.5 mM in 20 mM TRIS, pH 7.0) in the sitting drop; $Zn_4$ : RIDC–$2_4$ – 1 $\mu L$ of precipitant solution (100 mM BIS-TRIS, pH 6.5 , 25% PEG 3350, 3.4 mM $ZnCl_2$) and 2 μL protein (1.7 mM in 20 mM TRIS, pH 7.0) in the sitting drop; RIDC–$1_2$ – 1 μL of precipitant solution (26% PEG 2000) and 2 μL protein (3.8 mM in 20 mM TRIS, pH 7.0) in the sitting drop; $Cu_2$ : RIDC–$1_2$ – 1 $\mu L$ of precipitant solution (100 mM HEPES, pH 7.5, 200 mM NaCl, 25% PEG 3350, 10.7 mM $CuSO_4$) and 2 μL protein (5.4 mM in 20 mM TRIS, pH 7.0) in the drop. Appropriate crystals were transferred to a solution of mother liquor containing 20% glycerol as a cryoprotectant and frozen in liquid nitrogen or directly in the cryostream prior to measurement. X-ray diffraction data for $Zn_4$ : RIDC–$1_4$ and $Zn_4$ : RIDC–$2_4$ were collected at 100 K at the Stanford Synchrotron Radiation Laboratory (BL 9-2 and 7-1, respectively) using 1 – Å radiation. The data were integrated and scaled using Denzo/SCALA. Data for metal-free RICD–$1_2$ and $Cu_2$ : RIDC–$1_2$ were collected at 100 K using a Bruker Apex II CCD detector and monochromatized Cu-Kα radiation (1.54 Å) produced by a Siemens sealed tube source. The data were processed using SAINT and Bruker SADABS. All structures were determined through molecular replacement with MOLREP (4) by using the cyt $cb_{562}$ monomer structure (PDB ID:2BC5) as the search model. Rigid-body, positional, thermal and TLS refinement with REFMAC by using appropriate non-crystallographic symmetry restraints (5), along with manual rebuilding, and water placement with XFIT (6) produced the final models. All figures were produced with PYMOL (7). See Table S1 for x-ray data collection and refinement statistics.

**Protein Interface Redesign** For the cyt $cb_{562}$ (PDB ID: 2BC5) analysis, the following criteria and programs were used for defining designable interface clusters: AREAIMOL (8) was used to calculate SASA values and a residues with a SASA value below $10 n$Å2 were flagged as undesignable; LIGPLOT (9) was used to identify and flag as undesignable residues contacting the heme and Zn ligands in cyt $cb_{562}$ and $Zn_4$ : MBPC–$1_4$ (PDB ID: 2QLA) respectively; and WHATIF (10) was used to calculate the optimal hydrogen bond network, with residues making side chain-main chain hydrogen bonds scoring above 0.45 being flagged as undesignable.

All sequence design and rotamer optimizations were performed with fixed backbone templates, using a variant of the RosettaDesign algorithms (11) used for optimizing multiple conformers for a single sequence (12), in which each subunit of the oligomer represented a conformer. The variation from the previously published algorithm involves an improved rotamer optimization search algorithm in which: all rotamers are sampled in a random order; rotamers resulting in lower energies are duplicated in the rotamer set; and in each sampling iteration, the best of five independent simulated annealing trajectories are input into the next iteration where the rotamer set is reverted back to containing only unique rotamers. These improvements resulted in lower energies and increased sequence convergence with only a modest increase in sampling time.

For all rotamer optimizations, the Lennard–Jones van der Waals repulsive term was dampened and rotamers $\pm 1\sigma$ in torsion angles from canonical rotamers in $\chi 1$ for all amino acid types and in $\chi 2$ angles for aromatic amino acids were included in the rotamer set. For the native amino acid rotamer optimizations of $Zn_4$ : MBPC–$1_4$, four independent trajectories were performed and the resulting residue energies and SASAprob values (13) were averaged and ranked as described above. For redesigns, twenty trajectories were performed and the sequence of the lowest energy trajectory was selected for experimental characterization. Interface residues were defined as residues in which atoms of a residue on one chain were within 5.5 Å of atoms on a residue in another chain. Neighboring residues were defined in a similar manner.

**Docking Simulations** 3000 independent docking trajectories were carried out using RosettaDock (14).The two monomers that form the i1 interface in the $Zn_4$ : RIDC–$1_4$ design (the backbones of which are identical in $Zn_4$ : MBPC–$1_4$) were used as the starting structure for the docking simulations. One of the monomers was randomly spun along the axis connecting the centers of mass of both partners and the same monomer was also allowed to search a space of up to 3 Å normal to that axis, 8 Å in the plane perpendicular to the axis, and with up to an 8° tilt from the axis and an 8° additional spin around the axis. For docking simulations, these search parameters may be considered intermediary between a wide local search and a global search. Docking was performed with a full atom representation, in which rotamers were optimized and rotamers deviating slightly from canonical rotamers were sampled. Docking simulations were performed prior to solving the crystal structures of RIDC–$1_2$ and $Cu_2$ : RIDC–$1_2$.


1. Salgado EN, Faraone-Mennella J and Tezcan FA (2007) Controlling protein-protein interactions through metal coordination: Assembly of a 16-helix bundle protein. *J Am Chem Soc* 129:13374–13375.
2. Salgado EN, Lewis RA, Faraone-Mennella J, and Tezcan FA (2008) Metal-mediated self-assembly of protein superstructures: Influence of secondary interactions on protein oligomerization and aggregation. *J Am Chem Soc* 130:6082–6084.
3. Salgado EN, Lewis RA, Mossin S, Rheingold AL, and Tezcan FA (2009) Control of protein oligomerization symmetry by metal coordination: $C_2$ and $C_3$ symmetrical assemblies through $Cu^{II}$ and $Ni^{II}$ Coordination. *Inorg Chem* 48:2726–2728.
4. Vagin A and Teplyakov A (1998) MOLREP: An automated program for molecular replacement. *J Appl Cryst* 30:1022–1025.
5. Murshudov G, Vagin A, and Dodson E (1996) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst* D53:240–255.
6. McRee DE (1992) *J Mol Graphics* 10:44–46.
7. DeLano WL (2003) *The PYMOL Molecular Graphics System* (http://www.pymol.org).
8. Lee B and Richards FM (1971) Interpretation of protein structures—Estimation of static accessibility. *J Mol Biol* 55:379–400.

9. Wallace AC, Laskowski RA, and Thornton JM (1995) Ligplot—A program to generate schematic diagrams of protein ligand interactions. *Prot Eng* 8:127–134.
10. Vriend G (1990) What if—A molecular modeling and drug design program. *J Mol Graph* 8:52.
11. Liu Y and Kuhlman B (2006) RosettaDesign server for protein design. *Nucl Acids Res* 34:W235–238.
12. Ambroggio XI and Kuhlman B (2006) Computational design of a single amino acid sequence that can switch between two distinct protein folds. *J Am Chem Soc* 128:1154–1161.
13. Leaver-Fay A, Butterfoss G L, Snoeyink J, and Kuhlman B (2007) Maintaining solvent accessible surface area under rotamer substitution for protein design. *J Comp Chem* 28:1336–1341.
14. Gray JJ et al. (2003) Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *J Mol Biol* 331:281–299.

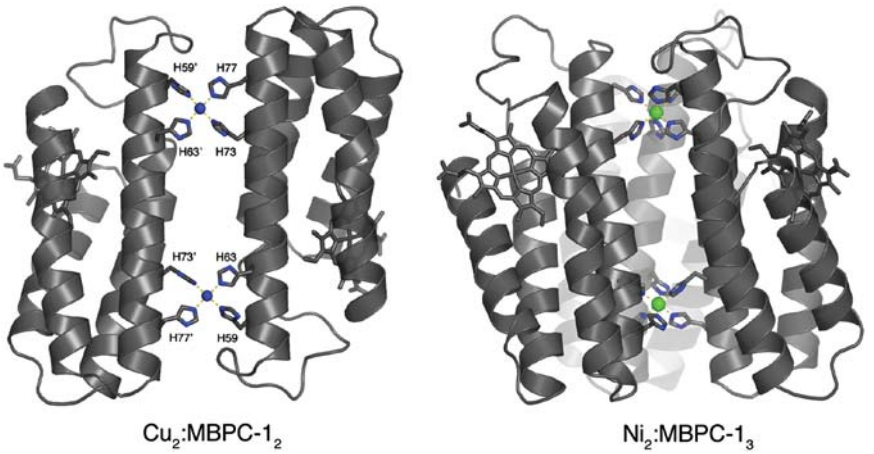


**Fig. S1.** Crystal structures of $C_2$ symmetrical $Cu_2$ : RIDC–$1_2$ and $C_3$ symmetrical $Ni_2$ : MBPC–$1_3$

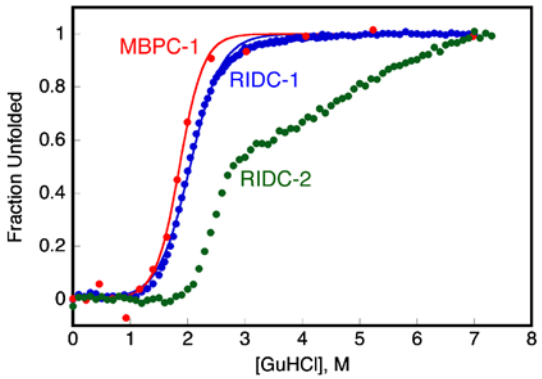


**Fig. S2.** Chemical denaturation profiles for MBPC-1, RIDC-1 and RIDC-2 monitored by CD spectroscopy (at 222 nm). The unfolding of MBPC-1 and RIDC-1 is well described by a two-state model, yielding free energies for unfolding of 5.2 and 4.3 kcal/M, respectively. The titrations were carried out using 5 μM protein in 20 mM TRIS buffer (pH 7) and 5 mM EDTA. Interestingly, RIDC-2 appears even more stable than both MBPC-1 and RIDC-1, judging from the onset of denaturation at higher guanidine hydrochloride (GuHCl) concentrations, although its unfolding does not appear to be two-state. We suggest that the additional hydrophobic residues present on the RIDC-2 surface lead to the formation of a folding intermediate not observed in the other two variants.

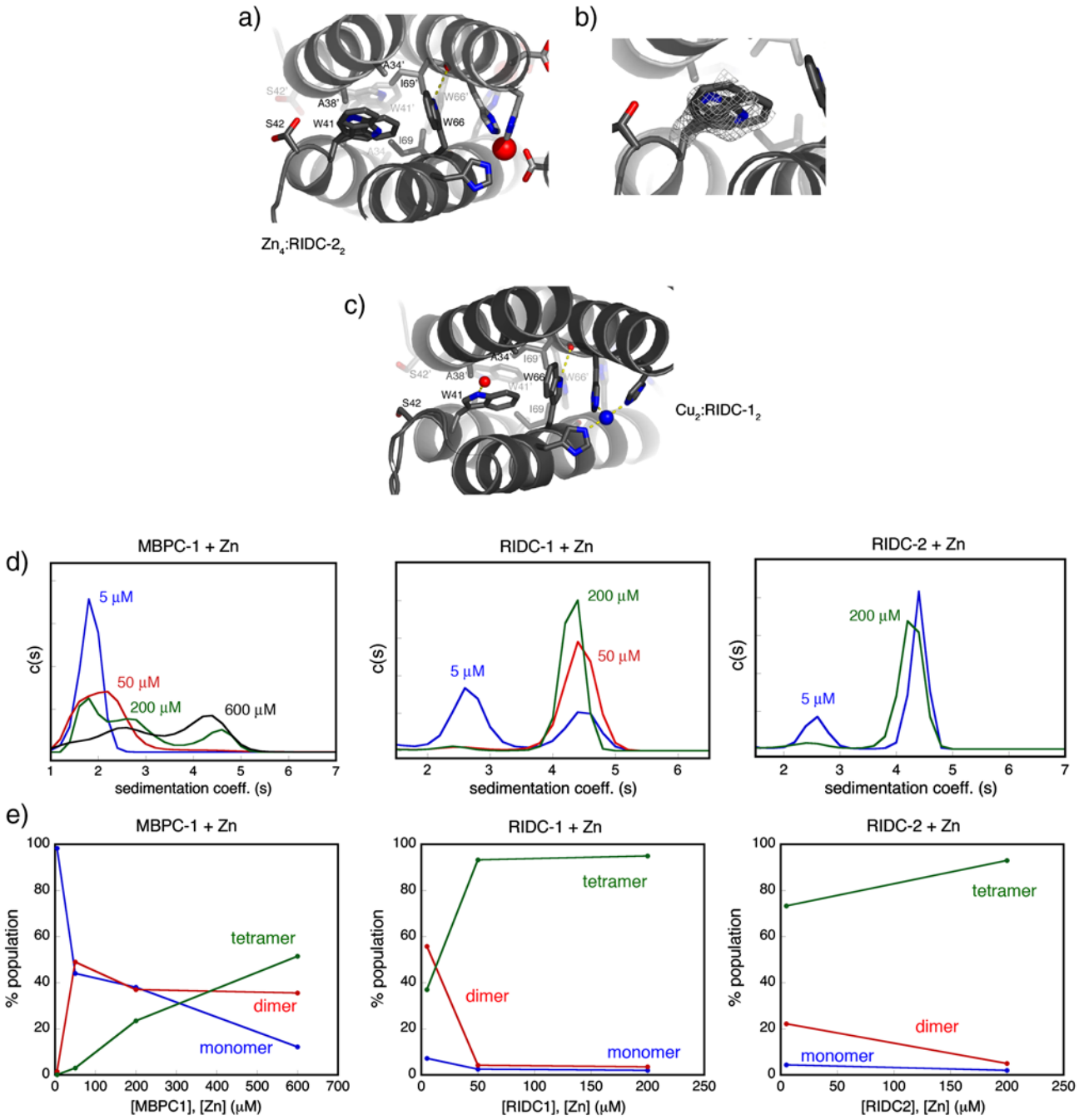


**Fig. S3.** i1 packing interactions in (a) $Zn_4$:RIDC-$2_2$ and (c) $Cu_2$ : RIDC–$1_2$. (b) Electron density showing the fluxionality of the Trp41 side chain in $Zn_4$ : RIDC–$2_2$. (d) Sedimentation coefficient distributions for various concentrations of MBPC-1, RIDC-1 and RIDC-2 in the presence of equimolar Zn(II). (e) Percent population of monomeric, dimeric and tetrameric species based on (d).

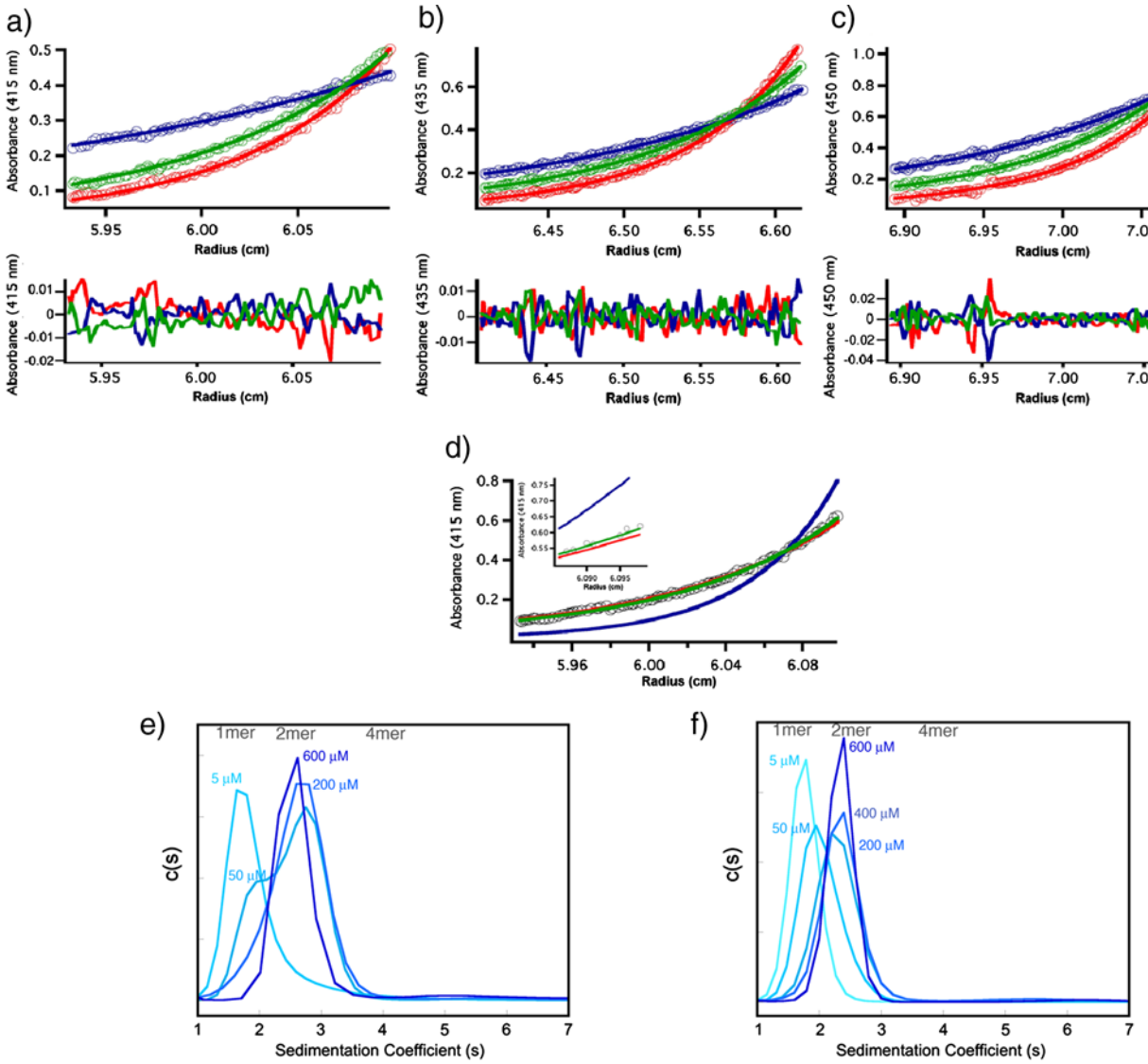


**Fig. S4.** Sedimentation equilibrium data for RIDC-1 in the presence of 5 mM EDTA. (a) 5 μM protein at 20,000, 30,000, and 35,000 rpm monitored at 41 nm. (b) 20 μM protein at 20,000, 25,000, and 30,000 rpm monitored at 435 nm. (c) 40 μM protein at 20,000, 25,000, and 30,000 rpm monitored at 450 nm. All samples were in 20 mM TRIS buffer (pH 7), and measurements were carried out at 25 °C. (d) Comparison of species analysis models used to globally fit SE data. Only data and fits for the 20,000-rpm scan of the 5 μM protein sample are shown (*Black Circles*), along with fits for the single species monomer (*Red*) and dimer (*Blue*), and monomer-dimer equilibrium (*Green*) models. (*Inset*) Close-up of the fits of the three different models. (e) Sedimentation coefficient distributions for various concentrations of RIDC-1 in the presence of 5 mM EDTA (shown in full scale). (f) Sedimentation coefficient distributions for various concentrations of RIDC-2 in the presence of 5 mM EDTA (shown in full scale).

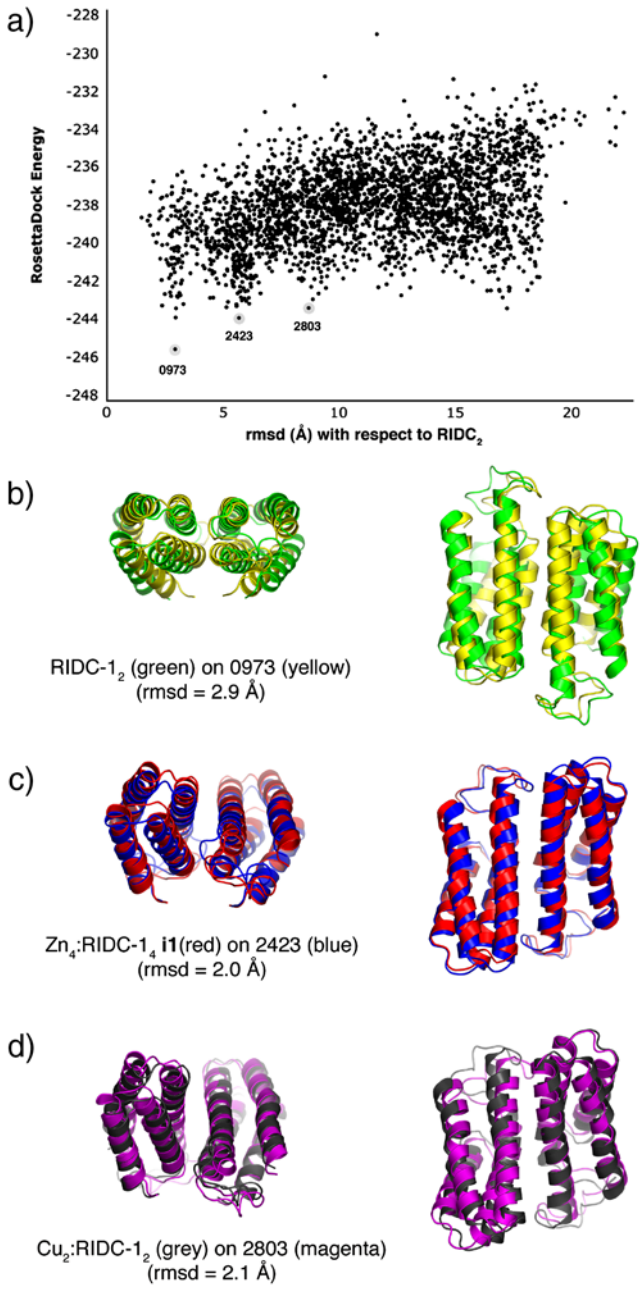


**Fig. S5.** Results from RosettaDock docking simulations (see also the corresponding Table S2). (a) Correlation between the docking energy and deviation from the RIDC–$1_2$ dimer crystal structure for 3,000 decoys. (b), (c) and (d) Superpositions of RIDC–$1_2$, $Zn_4$ : RIDC–$1_4$ and $Cu_2$ : RIDC–$1_2$ crystal structures with calculated decoy structures. The rmsd values shown in (a) are with respect to the RIDC – $1_2$ structures, whereas those in (b), (c) and (d) are for the shown superpositions.

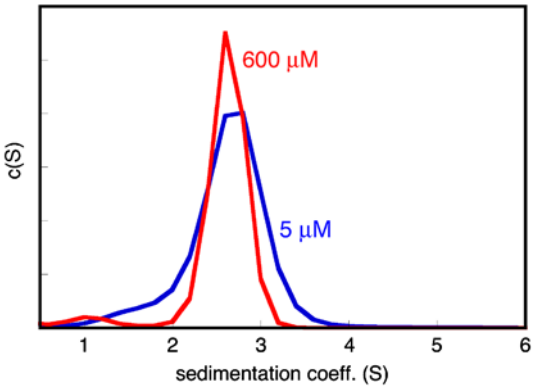


**Fig. S6.** Sedimentation coefficient distributions for 5 and 600 μM RIDC-1 in the presence of equimolar Cu(II), showing the exclusive formation of a dimeric species in solution.

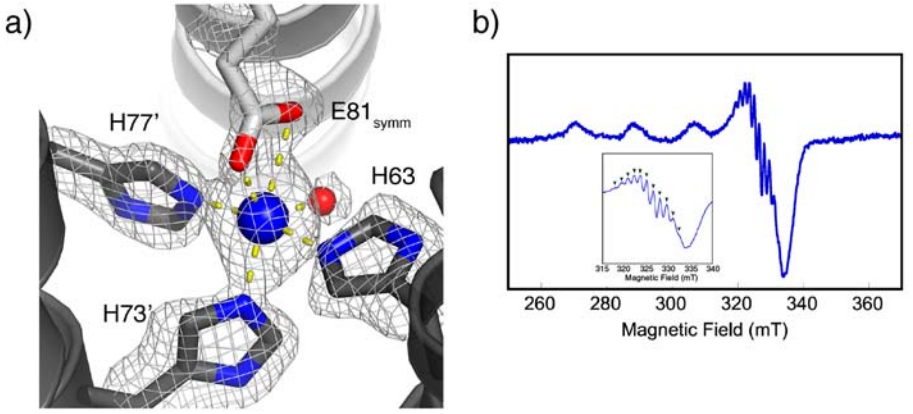


**Fig. S7.** (a) Cu coordination environment in $Cu_2$:RIDC–$1_2$ (Site 2), where the Glu81 side chain from a crystallographic symmetry related dimer is directly coordinated to Cu. (b) The X-band EPR spectrum of $Cu_2$:RIDC–$1_2$ (150 $\mu$MRIDC–1 + 100 μM Cu(II)) collected at 125 K. (inset) Close-up view of the g⊥ band, highlighting the superhyperfine splitting pattern. There are at least ten discernible superhyperfine peaks, whereas seven are expected for three equivalent nitrogen donors. We attribute this to a small degree of fluxionality and/or deviations from ideal square planar geometry in the Cu-coordination environment.

**Table S1. X-ray data collection and refinement statistics**

| | $Zn_4$:RIDC–$1_4$ | $Zn_4$:RIDC–$1_4$ | RIDC–$1_2$ | $Cu_2$:RIDC–$1_2$ |
|---|---|---|---|---|
| Location for data collection | SSRL BL 9-2 | SSRL BL 7-1 | UCSD | UCSD |
| Symmetry group | $P2_12_12_1$ | $P2_1$ | $P2_12_12_1$ | $C222_1$ |
| Unit cell dimensions | 57.2 Å × 87.9 Å × 152.2 Å | 48.4 Å × 65.2 Å × 70.8 Å, $\beta = 104.7^o$ | 49.3 Å × 55.5 Å × 72.4 Å | 66.2 Å × 87.0 Å × 80.8 Å |
| Resolution (Å) | 76.0 – 2.35 | 30.7 – 2.0 | 27.8 – 2.1 | 22.2 – 2.2 |
| X-ray wavelength (Å) | 0.979 | 0.976 | 1.542 | 1.542 |
| Number of unique reflections | 32072 | 26855 | 12023 | 11999 |
| Redundancy | 3.4 | 2.6 | 3.5 | 2.9 |
| Completeness (%)* | 98.3 (98.3) | 92.5 (92.4) | 98.7 (96.5) | 99.5 (98.4) |
| ⟨I / σI⟩* | 5.3 (1.4) | 11.0 (2.7) | 13.4 (4.1) | 6.5 (2.0) |
| ‡ (%)* | 11.8 (54.8) | 4.4 (17.6) | 7.8 (18.7) | 16.8 (48.2) |
| R§ (%)* | 22.6 (29.5) | 22.8 (27.1) | 20.0 (21.8) | 20.4 (23.3) |
| Free R¶ (%)* | 27.8 (35.8) | 27.2 (32.1) | 27.1 (20.6) | 27.0 (27.9) |
| Rmsd Bnd¶ (Å) | 0.011 | 0.009 | 0.007 | 0.008 |
| Rmsd Ang¶ (°) | 1.28 | 1.19 | 0.99 | 1.10 |
| Ramachandran plot (%) Most favored/Additionally allowed/ Generously allowed/Disallowed | 97.2/2.8/0.0/0.0 | 96.1/3.9/0.0/0.0 | 98.5/1.5/0.0/0.0 | 97.9/2.1/0.0/0.0 |

‡$R_{sym} = \Sigma\Sigma_j|I_j - \langle I\rangle|/\Sigma\Sigma_j|I_j|$.

§$R = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$.

‖Free R calculated against 7.2, 7.2. 6.9, and 6.9% of the reflections removed at random for the $Zn_4$:RIDC–$1_4$, $Zn_4$:RIDC–$2_4$, RIDC–$1_2$, and $Cu_2$:RIDC–$1_2$ structures, respectively.

¶Root mean square deviations from bond and angle restraints.

*Numbers in parentheses correspond to the highest resolution shell.

**Table S2. Results from RosettaDock docking simulations. A corresponding correlation between RosettaDock energies and rms deviations from the RIDC–$1_2$ structure is shown in Fig. S5**

| Superpositions* | | | |
|---|---|---|---|
| Structure 1 | Structure 2 | RMSD** | RosettaDock Energy |
| RIDC–$1_2$ | $Zn_4$ : RIDC–$1_4$ | 6.164 | N/A |
| RIDC–$1_2$ | 0973*** | 2.894 | -245.62 |
| RIDC–$1_2$ | 2423 | 5.644 | -243.98 |
| RIDC–$1_2$ | 2803 | 8.651 | -243.45 |
| $Zn_4$ : RIDC–$1_4$ | 0973*** | 5.384 | -245.62 |
| $Zn_4$ : RIDC–$1_4$ | 2423 | 1.998 | -243.98 |
| $Cu_2$ : RIDC–$1_2$ | 2803 | 2.090 | -243.45 |

*Crystal structures are indicated by bold font and docking simulation decoys are indicated by italic font. The decoys 0973, 2423, 2803 presented in this table are highlighted in Fig. S5, and respectively have energies 3.66, 2.87, and 2.61 standard deviations from the simulation mean.

**RMSD in Å over all α-Cs. In the case of tetrameric crystal structures, only the monomers forming the i1 interface are used in the superposition.

***0973.pdb is the lowest energy decoy of the simulation.