Published in final edited form as:
J Am Chem Soc. 2013 September 11; 135(36): . doi:10.1021/ja403503m.

Computational Design of an Unnatural Amino Acid Dependent Metalloprotein with Atomic Level Accuracy

Jeremy H. Mills[1], **Sagar D. Khare**[1,†], **Jill M. Bolduc**[2], **Farhad Forouhar**[3], **Vikram Khipple Mulligan**[1], **Scott Lew**[3], **Jayaraman Seetharaman**[3], **Liang Tong**[3], **Barry L. Stoddard**[2], and **David Baker**[1,4,5,*]

[1]Department of Biochemistry, University of Washington, Seattle, Washington, USA.

[2]Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA.

[3]Department of Biological Sciences, Northeast Structural Genomics Consortium, Columbia University, New York, New York, USA.

[4]Howard Hughes Medical Institute, University of Washington, Seattle, Washington, USA.

[5]Biomolecular Structure and Design Program, University of Washington, Seattle, Washington, USA.

Abstract

Genetically encoded unnatural amino acids could facilitate the design of proteins and enzymes of novel function, but correctly specifying sites of incorporation, and the identities and orientations of surrounding residues represents a formidable challenge. Computational design methods have been used to identify optimal locations for functional sites in proteins and design the surrounding residues, but have not incorporated unnatural amino acids in this process. We extended the Rosetta design methodology to design metalloproteins in which the amino acid (2,2'-bipyridin-5yl)alanine (Bpy-Ala) is a primary ligand of a bound metal ion. Following initial results that indicated the importance of buttressing the Bpy-Ala amino acid, we designed a buried metal binding site with octahedral coordination geometry consisting of Bpy-Ala, two protein based metal ligands, and two metal bound water molecules. Experimental characterization revealed a Bpy-Ala mediated metalloprotein with the ability to bind divalent cations including Co^{2+}, Zn^{2+}, Fe^{2+}, and Ni^{2+}, with a K_d for Zn^{2+} of ~40 pM. X-ray crystallographic analysis of the designed protein shows only slight deviation from the computationally designed model.

Introduction

Methods allowing site-specific incorporation of genetically encoded unnatural amino acids (UAAs) into proteins[1] offer great promise for the creation of biomolecules with novel function. Physical and chemical properties of proteins that emerged in biological systems only after eons of selective pressure can now be encoded in a single amino acid side chain.[2] Designing new functional sites with UAAs will require optimizing the chemical and structural environment around the UAA for the desired function. For example, a designed

*Corresponding Author dabaker@u.washington.edu.
†Present Addresses Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey, USA.

Supporting Information. Example theozymes, protein expression and characterization data, and detailed methods of computational design, protein expression, and K_d determinations. This material is available free of charge via the Internet at http://pubs.acs.org.

The structures of CB_02 and MB_07 bound to Co and Ni have been deposited in the RCSB Protein Databank under accession codes 4J9T, 4IWW, and 4IX0 respectively.

enzyme utilizing an UAA to provide a potent nucleophile or electrophile would also require appropriately positioned amino acids to correctly orient the UAA, to carry out additional catalytic roles (such as proton transfer), and to bind the substrate in an orientation suitable for catalysis. Computational protein design methods can, in principle, solve this problem by simultaneously positioning multiple residues, but such methods have focused thus far on design using only the 20 naturally occurring amino acids.

Metalloproteins carry out a diverse spectrum of important biological functions including oxygen transport, photosynthesis, respiration, nitrogen fixation, and water oxidation.[3] The ability to precisely design metal binding sites in proteins would in turn facilitate the design of new enzyme catalysts, as well as novel metal dependent protein folding or protein-protein interactions. Considerable work has aimed at designing novel metalloproteins,[4-8] but achieving atomic level accuracy has been challenging, as there are generally many alternative ways to form a metal binding site from flexible polar amino acid sidehcains.[9-11] Two UAAs with the ability to bind metal have been added to the genetic code of E. coli to date. Both (2,2'-bipyridin-5yl)alanine (Bpy-Ala)[12] and (8-hydroxyquinolin-3-yl)alanine (HQ-Ala)[13] have an inherent ability to bind diverse sets of metal ions with various affinities. From a computational protein design perspective, bidentate metal binding UAAs reduce the complexity of metal binding site design because of high inherent affinity for metal and by eliminating the need to make hydrogen bonds to unsatisfied donors or acceptors (as would be the case for the unliganded nitrogen in a His side chain contributing to a metal binding site).

Here, we describe the use of the Rosetta computational design methodology to engineer a metal binding site in a protein initially lacking that capacity utilizing a genetically encoded Bpy-Ala as the primary metal ligand. A first round of design provided insight into factors that must be considered when utilizing UAAs with strong intrinsic metal binding affinities in design. A second round of design resulted in the identification of a Bpy-Ala mediated metal binding site in which an octahedral coordination sphere around the bound metal ion is formed by Bpy-Ala, natural amino acid side chains, and explicitly modeled bound water molecules. The designed protein binds a range of divalent cations, including Zn^{2+} (with a measured dissociation constant in the pM range), Co^{2+}, Ni^{2+}, and Fe^{2+}, and the X-ray crystal structure is nearly identical to the design model.

Results

Lessons from initial design efforts

We first set out to design an active site to catalyze oxidative ring opening of catechol substrates. Four theoretical enzyme active sites ('theozymes', Supplementary Figure 1) were constructed through analysis of a crystal structure of a small molecule complex of bipyridine (Bpy), iron, and 3,6-di-tert-butylcatechol from the Cambridge Structural Database (CSD entry ZAZHAJ). As in natural catechol dioxygenases, the iron atom in this small molecule structure is coordinated by the phenols of the catechol moiety. Components of the theozymes included Bpy-Ala, the catechol containing small molecule dopamine, and two additional ligands to the metal that were allowed to be tyrosine or histidine. RosettaMatch[14] was used to identify sets of backbone positions in naturally occurring protein scaffolds with geometries that allow recapitulation of the theozyme geometries. For each of these "matches", RosettaDesign[15] was used to introduce additional interactions to stabilize the Bipy-Ala complex, and the resultant designs were filtered on how well the geometric constraints were satisfied by the design, whether or not the metal ligands were displaced during an attempt to repack designed side chains in the absence of the metal and dopamine ligand, and the orientation of the dopamine substrate within the active site. Genes encoding 13 designs were synthesized and the designed recombinant proteins were expressed in *E.*

coli in the presence and absence of the Bpy-Ala. Because Bpy-Ala incorporation is specified with the Amber stop codon, translation of the full-length protein should only occur in the presence of the UAA. Five of the designed proteins yielded soluble, full-length protein in the presence of Bpy-Ala, with little or no full-length protein observed in its absence (Supplementary Figure 2 and Supplementary Table 1), suggesting that expression of the proteins was dependent on the presence of Bpy-Ala.

Complexes of bipyridine exhibit π—π* transitions and metal—ligand charge-transfer (MLCT), both of which give rise to spectroscopic signatures in the UV or visible ranges.[16,17] Immediately after purification, the proteins were subjected to spectroscopic analysis in the range of 230 – 700 nm to probe π—π* (240 and 310 nm) or MLCT electronic transitions. MLCT wavelengths depend on the identity of the bound metal, but generally fall within the visible range.[16,17] Of the five proteins examined, two (CB_02 and CB_12) were observed to have a spectroscopic signature in a wavelength range indicative of a complex formed from Bpy-Ala and Fe^{2+} (Figure 1a and Supplementary Figure 3).

The crystal structure of CB_02, a design based on the scaffold with PDB ID 1eus, was solved to 1.4 Å resolution (Figure 1b). Unexpectedly, the loop on which the Bpy-Ala UAA resides was flipped out into solvent, and no electron density corresponding to either Bpy-Ala or the bound metal was observed. Bpy forms very stable homo-trimeric complexes with a number of divalent cations including Ni^{2+}, Co^{2+}, Fe^{2+}, and Zn^{2+}, and the spectroscopic data observed for CB_02 (Figure 1a) are consistent with the formation of a $[Fe(Bpy)_3]^{2+}$ complex (Figure 1a and Supplementary Figure 3). It is likely that the protein-bound Bpy-Ala forms a very stable $[Fe(Bpy\text{-}Ala_f)_2(Bpy\text{-}Ala_p)]^{2+}$ complex where $(Bpy\text{-}Ala_f)$ is free Bpy-Ala in the cell, and $(Bpy\text{-}Ala_p)$ is Bpy-Ala that has been incorporated in the protein. The formation of such a complex eliminates designed protein-metal interactions, and cannot be sterically accommodated by the scaffold, resulting in its exclusion into solvent. Because no density for Bpy-Ala was observed in the structure, the presence of Bpy-Ala in CB_02 was confirmed by mass spectrometric analysis (Supplementary Figure 4). Design CB_12 showed a similar spectroscopic signal to that observed for CB_02 (Figure 1a), and may also form a solvent exposed $[Fe(Bpy\text{-}Ala_f)_2(Bpy\text{-}Ala_p)]^{2+}$ complex.

The crystallographic information obtained for CB_02 suggested that Bpy-based design efforts should utilize negative design[18] – attempting to disfavor alternative, non-productive structures – to ensure that the Bpy-Ala remains constrained within the protein scaffold. With this in mind, a second round of computational design was undertaken.

Second-Round Design Calculations

Since the first round design results suggested that controlling the Bpy-Ala conformation was critical, we decided to focus primarily on control of metal binding site geometry in second round calculations. In order to simplify other aspects of the design problem, dopamine was replaced with two explicitly modeled water molecules. To better constrain the Bpy-Ala side chain within the protein, it was incorporated into elements of stable secondary structure during the RosettaMatch calculations. To increase the geometric agreement between the best matches and the theozymes, we included Asp and Glu as liganding residues in addition to His. This resulted in an increased number of matches, and hence the best matches have geometries closer to ideal. In order to favor binding of divalent cations that interact with Bpy in octahedral geometries, Tyr was eliminated, as this residue is often observed in trivalent cation binding sites. The Asp and Glu residues were placed so as to make hydrogen bonds to the bound water ligands. As described below, to further favor recapitulation of the sites as designed, we required that the metal binding histidine and bound water molecules form at least one hydrogen bond to the remainder of the protein; the Asp and Glu already

make two interactions in the theozymes-one to the metal and one to the water ligands. An example of a theozyme is shown in Figure 2.

In all, 15 octahedral coordination geometries were generated (Supplementary Figure 5) such that the bound hydroxyls were in axial-axial, mixed axial-equatorial, and equatorial-equatorial orientations, and contained Bpy-Ala, and Asp, Glu, or His residues in the remaining four coordination sites. Rosetta-Match was used to identify placements of the theozymes in native protein scaffolds, with the requirement that Bpy-Ala be placed on an element of secondary structure. The elimination of the requirement that the bulk of an exogenous ligand be accommodated in the designed proteins resulted in a greater number of initial matches than previously observed. For each of the theozyme placements, RosettaDesign was used to optimize the amino acid identities in the second shell to hold the metal binding residues in place *via* hydrogen bonds and packing interactions. For each design, unrestricted side chain repacking calculations were carried out, and designs in which the metal binding site geometry was completely preserved and each water and histidine ligand made at least one hydrogen bond were selected. To reduce possible problems in folding, designed residues beyond the first shell (directly interacting with the metal) and second shell (interacting with first shell residues) were reverted to their native identities, and genes encoding 28 designs were synthesized. The numbers of matches and designs that satisfied the aforementioned criteria are listed in Supplementary Figure 6.

Of the 28 designs, 9 expressed solubly in *E. coli* only when Bpy-Ala was included in the expression medium, and two showed expression in both the presence and absence of Bpy-Ala (Supplementary Figure 7). The 9 purified proteins showing clear Bpy-Ala dependent expression were subjected to spectroscopic analysis in the range of 230-650 nm. Four of the proteins tested exhibited a shoulder on the tryptophan A_{280} peak from ~290-330 nm consistent with the π—π * transitions of Bpy-Ala metal complexes (Figure 3a), and only one protein showed signal in the range of 450-600 nm where $[Fe(Bpy)_3]^{2+}$ complexes absorb (Supplementary Figure 8a). Thus the Bpy-Ala residues in all but one of the second-round designs that expressed appear to be sufficiently buried that formation of the $[Fe(Bpy\text{-}Ala_f)_2(Bpy\text{-}Ala_p)]^{2+}$ complex is disfavored.

The design with the most pronounced absorbance in the range of 290-330 nm, MB_07 (Figure 3a), utilizes Bpy-Ala, Asp, Glu, and two bound waters as metal ligands. Of the remaining designs with signal in the range of 290-330 nm, only MB_02 utilized His as one of the metal ligands. Hence, MB_02 and MB_07 were subjected to crystallographic analysis.

Crystal Structures of MB_07 Bound to Co^{2+} and Ni^{2+}

The structure of design MB_07 (based on PDB 1igs) was solved to 2.3 Å resolution. The overall topology of the starting scaffold was maintained in the solved structure (backbone r.m.s.d = 0.80 Å). In the designed metal binding site, electron density corresponding to Bpy-Ala, a bound metal ion, and a single water molecule was clearly observed (Figure 4a). As in the design model, the metal ion was coordinated by E159, D184 and Bpy-Ala. While both acidic residues make monodentate interactions with the metal in the design model, E159 coordinates the metal in a bidentate fashion and thus displaces one of the metal bound waters in the design (Figs. 4a and 4b). The single metal bound water forms a hydrogen bond with residue S180, an interaction that was explicitly defined in the design. The all-atom rmsd of all residues within 6 Å of the bound metal was only 0.9 Å (Figure 4b). Although Bpy-Ala was modeled as being completely planar, a slight deviation from planarity was observed in the solved structure (Figure 4a and b).

X-ray fluorescence analysis was carried out at the Advanced Light Source (Lawrence Berkeley National Laboratory), and the identity of the bound metal was determined to be

cobalt (Supplementary Figure 9a). This is likely due to initial purification of the designs on TALON resin (Clonetech, Mountain View, CA), which is an immobilized metal affinity purification resin that contains Co^{2+}; bipyridine forms octahedral complexes with Co^{2+}.[19] To determine if MB_07 has affinity for other metals that interact with bipyridine in an octahedral fashion, purifications of MB_07 on affinity resins loaded with Ni^{2+} and Zn^{2+} were carried out. MB_07 purified on both Ni^{2+} and Zn^{2+} had similar, but not identical absorbance spectra to that after purification on Co^{2+} resin (Figure 3b). To examine whether the binding site geometry depends on the bound metal, the Ni^{2+} and Zn^{2+} loaded proteins were also subjected to crystallographic analysis. The structure of Ni^{2+} bound MB_07 was solved to 2.5 Å resolution, and the presence of Ni^{2+} in the protein was again confirmed with X-ray fluorescence (Supplementary Figure 9b). Overall, the binding mode was identical to that observed in the Co^{2+} loaded protein (Figure 4c and d) except that Bpy-Ala in the Ni^{2+} bound protein appeared to deviate less from planarity relative to the Co^{2+} bound structure. The all atom rmsd of residues within 6 Å of the bound metal was 1.0 Å in the Ni^{2+} bound structure. Difference density corresponding to a water molecule was observed in the Ni^{2+} bound structure at the same position as in the Co^{2+} structure in an F_o-F_c map (calculated after a round of refinement in REFMAC in the absence of modeled water) (Supplementary Figure 10), suggesting a metal bound water molecule is present in the Ni^{2+} structure as well. MB_07 also binds Fe^{2+} (Supplementary Figure 11), but we were unable to solve the structure of that complex.

Metal Binding Affinity of MB_07

To determine the affinity of MB_07 for metal, we first titrated Zn^{2+} into apo MB_07 (see supplementary methods), and monitored the change in absorbance at 310 nm (Supplementary Figure 12a). The absorbance increased linearly until the metal and protein concentrations were equal, at which point no further increase in signal was observed. A non-linear least squares fit of these data suggested an upper limit of the K_d of 10 nM (Supplementary Figures 12a and b).

To increase sensitivity at low Zn^{2+} concentrations, we used the Zn^{2+} sensitive chromophore 4-(2-pyridylazo)resorcinol (PAR), which binds Zn with a Kd of 250 nM.[20] Removal of Zn from PAR was monitored through the change in absorbance at 500 nm as apo MB_07 was titrated into a solution of PAR and Zn^{2+}. As the concentration of apo MB_07 was increased, a linear decrease in absorbance at 500 nm was observed until the protein and metal concentrations were equal, at which point the A_{500} was indistinguishable from apo PAR (Supplementary Figure 12c). Fits of these data suggest the K_d of MB_07 for Zn^{+2} is less than 1 nM (Supplementary Figures 12c and d).

To probe still lower Zn^{2+} concentrations, the Zn^{2+} sensitive fluorescent probe FluoZin-3 (Life Technologies), with a K_d for Zn^{2+} of 15 nM,[21] was used to carry out a competition assay similar to that described with PAR. A fit of the fluorescence intensity as a function of protein concentration yielded an estimate of the K_d of MB_07 for Zn^{2+} of 37 ± 15 pM (Figure 5 and Supplementary Figure 14). Because the K_d value obtained from the fits is far lower than the K_d of FluoZin-3 for Zn^{2+}, it was difficult to measure this value precisely, and 37pM is likely an upper bound.

Discussion

The combination of genetically encoded UAAs with computational protein design methods allows control over both the placement of the UAA within a protein, and the interactions between the UAA and the surrounding protein environment. One reason that metalloprotein design has been difficult is that the precise orientations of multiple protein side chains must be explicitly defined. Not only is the entropic cost of orienting the two pyridine rings of

Bpy-Ala in a geometry compatible with metal binding paid during the synthesis of the small molecule itself, it is also uncharged, and therefore lends itself to efforts focused on design of more buried metal binding sites. Furthermore, as observed in the structure of MB_07, the orientation of Bpy-Ala can be buttressed through hydrophobic packing interactions, which is not the case in metal binding sites constructed from polar residues, which are often stabilized through extensive hydrogen bonding networks.

Our initial designs with Bpy-Ala on portions of the protein not well constrained by the remainder of the scaffold resulted in the extrusion of Bpy-Ala to solvent, likely due to the formation of a $[Fe(Bpy\text{-}Ala_f)_2(Bpy\text{-}Ala_p)]^{2+}$ complex. Modifications to the initial design protocol were implemented to address this problem, and a second round of design was carried out. Of nine Bpy-Ala containing second round designs that expressed solubly, only one exhibited a spectroscopic signal indicative of formation of a $[Fe(Bpy\text{-}Ala_f)_2(Bpy\text{-}Ala_p)]^{2+}$ complex (Supplementary Figure 8a). Structural analysis of the second round design MB_07, confirmed the formation of an octahedral metal binding site as designed. Not only is Bpy-Ala relatively buried in this design, it is sandwiched between leucine and phenylalanine residues from the native scaffold. These data suggest that metal binding UAA cofactors should be well constrained within a protein scaffold to realize the desired functional site geometry.

Competition experiments suggest the affinity of MB_07 for Zn^{2+} is approximately 40 pM. The K_d of Bpy for Zn^{2+} is reported to be 7.5 μM,[19] hence the additional designed residues contribute substantially to affinity.

MB_07 was designed to have metal binding ability, but not enzymatic activity. In certain metalloenzymes (e.g. zinc hydrolases), bound metal ions act as Lewis acids that serve to activate substrates for nucleophilic attack.[22] Moreover, a metal bound water molecule, itself activated by the metal, often serves as the active site nucleophile in such cases.[22] While a close match was observed between the designed metal binding site and the solved structure, higher accuracy in the coordination geometry would be required for success with a designed metalloenzyme. Nonetheless, this designed metal binding site could provide the starting point for the subsequent design of UAA dependent hydrolases. More generally, the methods developed and insights gained in this study should be directly useful for designing a wide range of UAA based functional sites.

Conclusion

We report the extension of computational protein design methods to generate a protein with the ability to bind metals with high affinity by exploiting the inherent chemical functionality of a genetically encoded unnatural amino acid. Structural characterization revealed the importance of controlling the conformation of the introduced unnatural amino acid. The methods and insights from this work will inform the design of unnatural amino acid based enzyme catalysts.

Methods

Protein expression, purification, and analysis

Designed proteins were ordered from Genscript in a pET29b expression plasmid (EMD Millipore), and contained an Amber stop codon at the desired position of Bpy-Ala incorporation. Protein expression was carried out in the *Escherichia coli* BL21(DE3) cell line. BL21(DE3) cells were co-transformed with each design in pET29b and pEVOL-BpyRS (Schultz Laboratory, The Scripps Research Institute).[23] The pEVOL-BpyRS plasmid contains an orthogonal tRNA with an anticodon loop specific to the Amber stop codon, and

an aminoacyl tRNA synthetase evolved to acylate the orthogonal tRNA with Bpy-Ala. Protein expression was carried out in Terrific Broth at 18 °C in the presence of 500 μM *rac*-Bpy-Ala. Protein purification was carried out using TALON resin (Clonetech), Ni-NTA resin (Qiagen), or Zn^{+2} loaded Hi-trap chelating resin (G.E. Biosciences). Protein purity and size was analyzed using SDS-PAGE analysis. All absorbance spectra were collected on a Molecular Devices Spectramax M5^{e} microplate reader.

CB_02 structure determination

CB_02 was purified on TALON resin, followed by anion exchange chromatography and gel filtration. Purified protein was concentrated to 5 mg/mL in 10 mM Tris buffer, pH 8.0. Crystals were grown using the micro-batch method at 21 °C and a precipitant solution of 100 mM sodium cacodylate (p.H. 6.5) and 15% (w/v) PEG 3350. Solutions were supplemented with 5 mM $FeCl_2$, and crystals were grown in a COY anaerobic glove box with O_2 levels maintained below 2 ppm to prevent oxidation of the iron. Crystals were flash frozen in liquid nitrogen after immersion in precipitant solution supplemented with 25% (v/v) glycerol, and data were collected at 100 K. A native diffraction data set to resolution 1.4 Å was collected on a single crystal of CB_02 at the X4C beamline of the National Synchrotron Light Source (NSLS). The diffraction images were processed with the HKL package.[24] The data processing statistics are summarized in Supplementary Table 3. The structure was determined by the molecular replacement method with the program COMO[25], using the structure of the bacterial sialidase (PDB id: 1EUU) as the search model. The complete atomic model was built with the program XtalView[26] and refined with the program CNS[27] and Phenix.[28] Since the side chain of the Bpy-Ala at position 92 is disordered in the structure, it was modeled as an alanine residue. The Cα r.m.s.d. between the design and solved structure was calculated to be 1.07 Å. The structure of CB_02 represents target OR61 of the North East Structural Genomics consortium.

MB_07 structure determination

Cobalt-bound MB_07 was purified on TALON resin followed by anion exchange and gel filtration. Purified protein was concentrated to 10 mg / mL in 10 mM Tris pH 7.6, with 50 mM NaCl. Crystals of cobalt-bound MB_07 grew in 1.8 M ammonium sulfate, 5% PEG 3350, 50 mM BisTris pH 6.5. Nickel-bound MB_07 was first purified on Ni-NTA resin (Qiagen), followed by anion exchange and gel filtration. Purified protein was again concentrated to 10 mg / mL in the buffer described above. Crystals grew in 1.6 M ammonium sulfate supplemented with 3% PEG 3350 and 50 mM BisTris at pH 6.5. In both cases, crystals were flash cooled after soaking in two stages of glucose cryoprotectant. Data were collected on an RaxisIV++. HKL2000[24] was used to process and scale the data to 2.3Å resolution for the MB_07-Co structure, and 2.5Å for the MB_07-Ni structure. CCP4[29] was used to determine phases by molecular replacement, using the structure with PDB ID 1igs with truncated design residues as a search model. Coot[30] and Refmac[31] were used for model building and refinement, respectively, excluding a random 5% of the data for cross validation. Statistics are shown in Supplementary Table 4. Identities of the Bipy bound metals were confirmed via X-ray fluorescence data collected at beamline 5.0.2 at the ALS (Advanced Light Source, Lawrence Berkely Laboratory, Berkeley, CA).

Unbiased density was observed for the unnatural amino acid Bpy-Ala with bound metal, which was the strongest feature in density, at 7.8 σ in the case of cobalt, and 8.6 σ for nickel. Co bound MB_07 showed a Cα r.m.s.d. from the design structure of 0.79 Å, while Ni bound MB_07 had a Cα r.m.s.d. 0.49 Å.

Identification of bound metals

X-ray fluorescence data were collected at the Advanced Light Source (ALS) synchrotron facility at the Lawrence Berkeley National Laboratory (University of California) on beamline 5.0.2 with the assistance of ALS staff. Metal edges of Cu, Ni, Co, Fe, Zn, and Mg were scanned using in house BOS software. Metal edge fluorescence data are shown in Supplementary Figure 9.

Determination of the K_d of MB_07 for Zn^{2+}

A competition assay was carried out in which MB_07 was titrated into a mixture of FluoZin-3 and $ZnSO_4$. The FluoZin-3 and $ZnSO_4$ concentrations were 1 μM, and 100 nM respectively, and apo MB_07 was present at concentrations ranging from 2 nM to 1 μM (see Supplementary Methods). All data were collected on a SpectraMax M5^e microplate reader in cuvette mode. Non-linear least squares fitting of the data was carried out using Origin 7.0 software.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Peter Schultz for the generous gift of the pEVOL-BpyRS plasmid, and Lubica Supekova for assistance in transferring the plasmid to us. We thank Christine Tinberg, Neil King, and Yifan Song for helpful discussions. We thank Randy Abramowitz and John Schwanof for setting up the X4C beamline, and Gaetano Montelione and Greg Kornhaber for assistance in facilitating the crystallographic analysis of CB_02. This research was supported by the Defense Advanced Research Projects Agency (HR001-08-0085 ARPA), and the Defense Threat Reduction Agency (HDTRA1-11-1-0041). Research reported in this publication was supported by NIGMS of the National Institutes of Health under award number F32GM099210 to J.H.M. The content is solely the responsibility of the authors and does not represent the official views of the National Institutes of Health.

REFERENCES

1. Liu CC, Schultz PG. Annu. Rev. Biochem. 2010; 79:413. [PubMed: 20307192]
2. Wang J, Xie J, Schultz PG. J. Am. Chem. Soc. 2006; 128:8738. [PubMed: 16819861]
3. Lippard, SJ.; Berg, JM. Principles of Bioinorganic Chemistry. University Science Books; Mill Valley: 1994.
4. Lu Y, Yeung N, Sieracki N, Marshall NM. Nature. 2009; 460:855. [PubMed: 19675646]
5. Lu Y. Curr. Opin. Chem. Biol. 2005; 9:118. [PubMed: 15811795]
6. DeGrado WF, Summa CM, Pavone V, Nastri F, Lombardi A. Annu. Rev. Biochem. 1999; 68:779. [PubMed: 10872466]
7. Hellinga HW. Fold. Des. 1998; 3:R1. [PubMed: 9502313]
8. Iverson BL, Iverson SA, Roberts VA, Getzoff ED, Tainer JA, Benkovic SJ, Lerner RA. Science. 1990; 249:659. [PubMed: 2116666]
9. Telmer PG, Shilton BH. J. Mol. Biol. 2005; 354:829. [PubMed: 16288781]
10. Der BS, Machius M, Miley MJ, Mills JL, Szyperski T, Kuhlman B. J. Am. Chem. Soc. 2012; 134:375. [PubMed: 22092237]
11. Salgado EN, Faraone-Mennella J, Tezcan FA. J. Am. Chem. Soc. 2007; 129:13374. [PubMed: 17929927]
12. Xie J, Liu W, Schultz PG. Angew. Chem. Int. Ed. Engl. 2007; 46:9239. [PubMed: 17893898]
13. Lee HS, Spraggon G, Schultz PG, Wang F. J. Am. Chem. Soc. 2007; 131:2481. [PubMed: 19193005]
14. Zanghellini A, Jiang L, Wollacot AM, Cheng G, Meiler J, Althoff EA, Röthlisberger D, Baker D. Protein. Sci. 2006; 15:2785. [PubMed: 17132862]

15. Kuhlman B, Baker D. Proc. Natl. Acad. Sci. USA. 2000; 97:10383. [PubMed: 10984534]
16. Meyer TJ. Pure & Appl. Chem. 1986; 58:1193.
17. Mason SF. Inorg. Chim. Acta. Rev. 1968; 2:89.
18. Hect MH, Richardson JS, Richardson DC, Ogden RC. Science. 1990; 249:884. [PubMed: 2392678]
19. Smith, RM.; Martell, AE. Critical Stability Constants, Vol. 2, Amines. Plenum Press; New York: 1975. p. 235-236.
20. Hunt JB, Neece SH, Ginsburg A. Anal. Biochem. 1985; 146:150. [PubMed: 3887984]
21. Gee KR, Zhou ZL, Qian WJ, Kennedy R. J. Am. Chem. Soc. 2002; 124:776. [PubMed: 11817952]
22. Matthews BW. Acc. Chem. Res. 1998; 21:333.
23. Young TS, Ahmad I, Yin JA, Schultz PG. J. Mol. Biol. 2010; 395:367.
24. Otwinowski Z, Minor W. Method. Enzymol. 1997; 276:307.
25. Jogl G, Tao X, Xu Y, Tong L. Acta. Cryst. D. 2001; 57:1127. [PubMed: 11468396]
26. McRee DE. J. Struct. Biol. 1999; 125:156. [PubMed: 10222271]
27. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang J-S, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Acta Cryst. D. 1998; 54:905. [PubMed: 9757107]
28. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. Acta Cryst. D. 2010; 66:213. [PubMed: 20124702]
29. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissnel EB, Leslie AGW, McCoy A, McNicholas GN, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. Acta Cryst. D. 2011; 67:235. [PubMed: 21460441]
30. Emsley P, Cowtan K. Acta Cryst. D. 2004; 60:2126. [PubMed: 15572765]
31. Vagin AA, Steiner RS, Lebedev AA, Potterton L, McNicholas S, Long F, Murshudov GN. Acta Cryst. D. 2004; 60:2284.

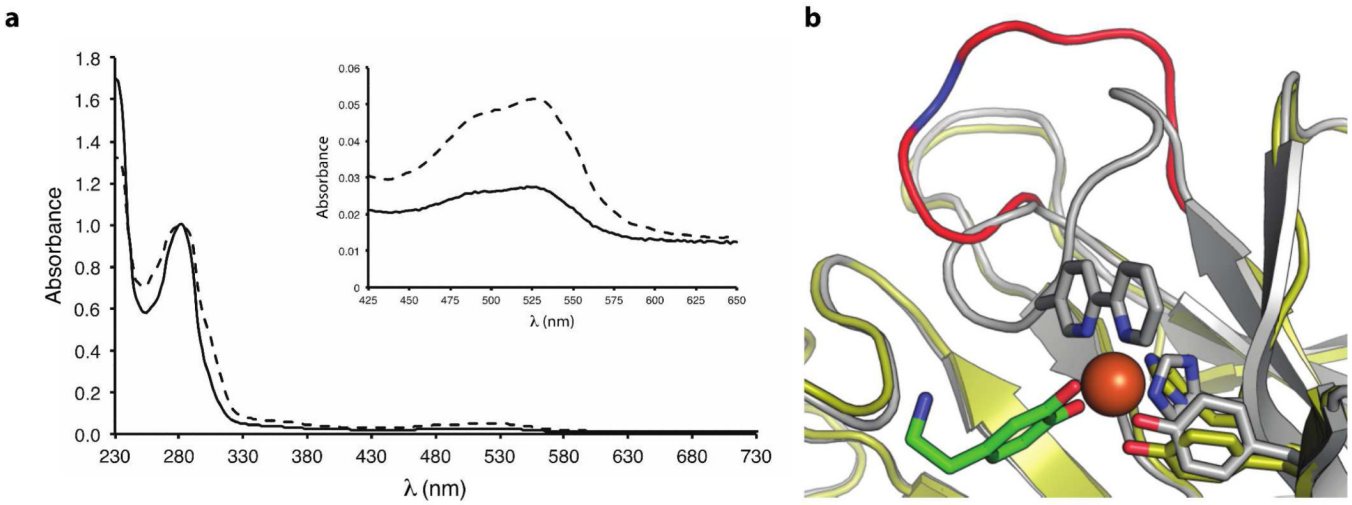

Figure 1.

Spectral and structural analysis of round 1 designs. (a) Absorbance spectra of designs CB_02 (solid line) and CB_12 (dashed line). The inset is an enlarged spectral region from 425-650 nm. The spectral signature in this region is consistent with a $[Fe(Bpy)_3]^{2+}$ complex. Spectra were normalized to an A_{280} of 1.0. (b) Superposition of CB_02 design with the structure solved at 1.4 Å. The design is shown in grey, the structure in yellow, and the dopamine ligand in green sticks. The Bpy-Ala containing loop in the CB_02 structure is colored red. No density corresponding to the Bpy-Ala side chain was observed in the structure, and the position of incorporation in the flipped out loop is colored blue.

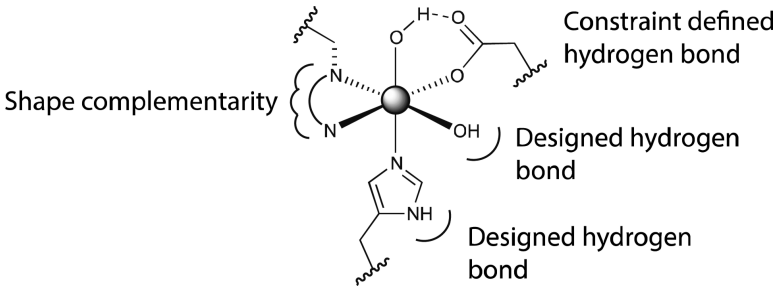

Figure 2.

An example theozyme for second-round design calculations. When Asp or Glu was included as a ligand of the bound metal (grey sphere), a hydrogen bond between a metal bound hydroxyl and a liganding acidic residue was specified as shown during the calculations. Additional interactions considered during design (hydrogen bonds to bound water molecules and metal bound histidines, and shape complementarity to Bpy-Ala) are shown as semicircles.

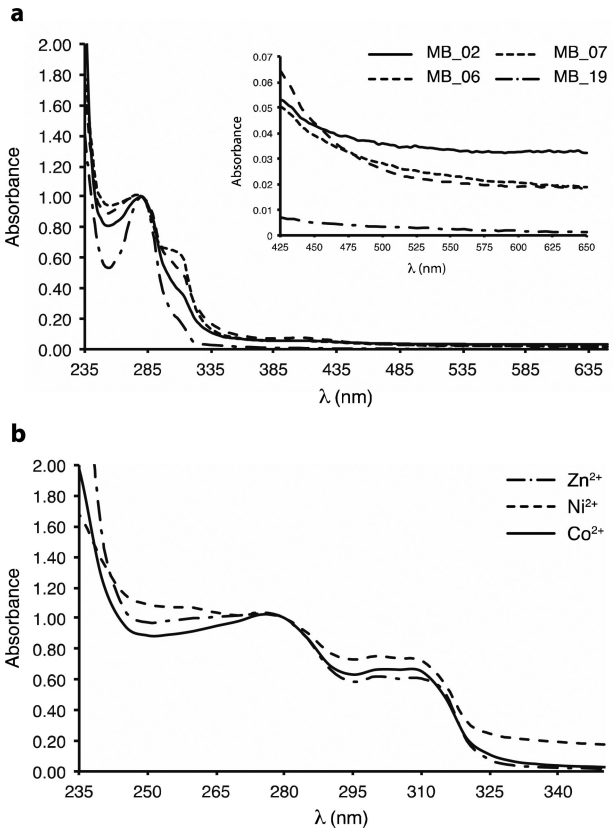

Figure 3.

Spectroscopic analysis of metal binding designs. (a) Round 2 designs MB_02, MB_06, MB_07, and MB_19 were analyzed spectroscopically from 230-650 nm. The presence of a shoulder in the range of 290-330 nm suggests the presence of Bpy bound metal ions. The absence of spectroscopic signal in the range of 450-600 nm (enlarged in the inset) indicates that a $[Fe(Bpy\text{-}Ala_f)_2(Bpy\text{-}Al_p)]^{2+}$ complex was not formed. (**b**) Absorbance spectra of MB_07 bound to Zn^{2+}, Ni^{2+} and Co^{2+} are shown. In spite of slight differences, all spectra contain signal in the range of 290-330 nm. Spectra are normalized to an A_{280} of 1.0.

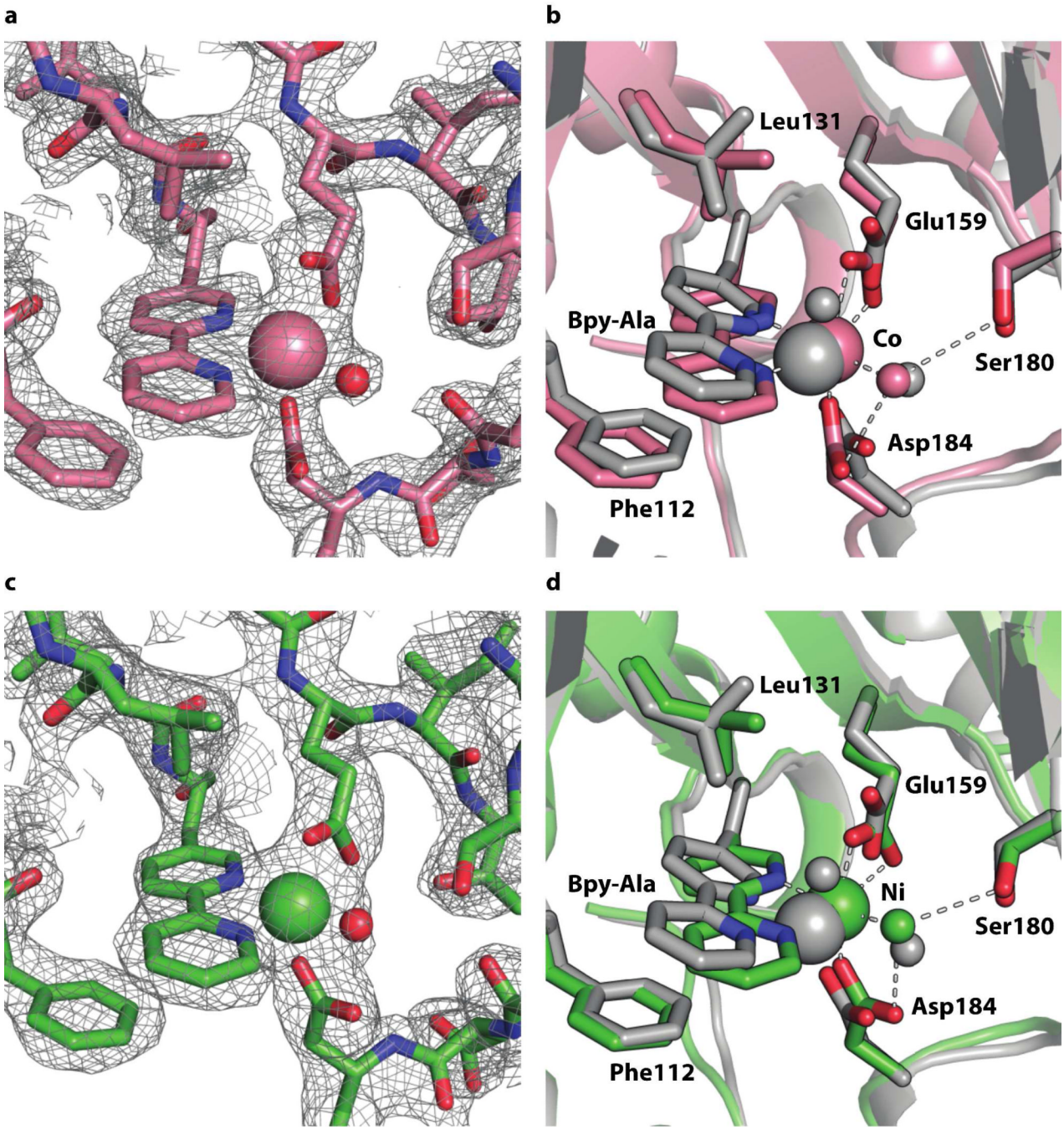

Figure 4.

X-ray crystallographic analysis of MB_07 bound to Co^{2+} and Ni^{2+}. Electron density from a $2F_o$-F_c map in the vicinity of the Bpy-Ala bound to Co^{2+} (**a**) and Ni^{2+} (**c**) is shown and is contoured at 1.0 σ. Density for Bpy-Ala, Co^{2+} or Ni^{2+} (pink and green spheres respectively), D184, E159, and a metal bound water molecule (red spheres) is visible. A comparison of the MB_07 design model (grey) with the solved crystal structure of MB_07 bound to Co^{2+} (**b**) and Ni^{2+} (**d**) is shown. The crystal structures of MB_07 bound to Co^{2+} and Ni^{2+} are shown in pink and green respectively. In all cases, bound metals are shown in large spheres and

bound waters are shown as small spheres. Dashed lines represent polar interactions observed in the crystal structures that are involved in metal coordination.

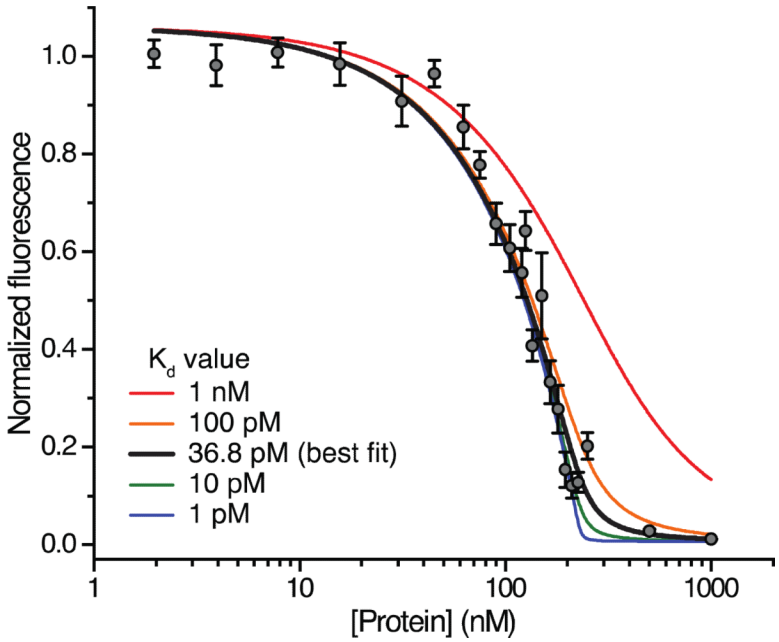

Figure 5.

Determination of MB_07 binding affinity. Normalized fluorescence values, proportionate to the concentration of FluoZin-3 bound to Zn^{2+}, are shown for $ZnSO_4$ and FluoZin-3 in the presence of apo MB_07 ranging in concentration from 2 nM to 1 μM (see supplementary information). Simulated fits were generated using equations 12-15 in the supplementary information, and are shown for K_d values ranging from 1 pM to 1 nM. The curve of best fit, corresponding to a K_d of 36.8 pM is shown in black.