Inorg. Chem. **2006**, *45*, 9941–9958

Forum

Femtomolar Zn(II) Affinity in a Peptide-Based Ligand Designed To Model Thiolate-Rich Metalloprotein Active Sites

Amy K. Petros,[†] Amit R. Reddi,[†] Michelle L. Kennedy,[†] Alison G. Hyslop,[‡] and Brian R. Gibney*[,†]

Department of Chemistry, Columbia University, 3000 Broadway, MC 3121, New York, New York 10027, and Department of Chemistry, St. John's University, 8000 Utopia Parkway, Queens, New York 11439

Received December 25, 2005

Metal–ligand interactions are critical components of metalloprotein assembly, folding, stability, electrochemistry, and catalytic function. Research over the past 3 decades on the interaction of metals with peptide and protein ligands has progressed from the characterization of amino acid–metal and polypeptide–metal complexes to the design of folded protein scaffolds containing multiple metal cofactors. De novo metalloprotein design has emerged as a valuable tool both for the modular synthesis of these complex metalloproteins and for revealing the fundamental tenets of metalloprotein structure–function relationships. Our research has focused on using the coordination chemistry of de novo designed metalloproteins to probe the interactions of metal cofactors with protein ligands relevant to biological phenomena. Herein, we present a detailed thermodynamic analysis of Fe(II), Co(II), Zn(II), and $[4Fe\text{-}4S]^{2+}$ binding to **IGA**, a 16 amino acid peptide ligand containing four cysteine residues, H_2N–KL**C**EGG–**C**IG**C**GA**C**–GGW–$CONH_2$. These studies were conducted to delineate the inherent metal-ion preferences of this unfolded tetrathiolate peptide ligand as well as to evaluate the role of the solution pH on metal–peptide complex speciation. The $[4Fe–4S]^{2+/+}$-**IGA** complex is both an excellent peptide-based synthetic analogue for natural ferredoxins and is flexible enough to accommodate mononuclear metal-ion binding. Incorporation of a single ferrous ion provides the Fe^{II}-**IGA** complex, a spectroscopic model of a reduced rubredoxin active site that possesses limited stability in aqueous buffers. As expected based on the Irving–Williams series and hard–soft acid–base theory, the Co(II) and Zn(II) complexes of **IGA** are significantly more stable than the Fe(II) complex. Direct proton competition experiments, coupled with determinations of the conditional dissociation constants over a range of pH values, fully define the thermodynamic stabilities and speciation of each M^{II}-**IGA** complex. The data demonstrate that Fe^{II}-**IGA** and Co^{II}-**IGA** have formation constant values of 5.0×10^8 and 4.2×10^{11} M^{-1}, which are highly attenuated at physiological pH values. The data also evince that the formation constant for Zn^{II}-**IGA** is 8.0×10^{15} M^{-1}, a value that exceeds the tightest natural protein Zn(II)-binding affinities. The formation constant demonstrates that the metal–ligand binding energy of a $Zn^{II}(S–Cys)_4$ site can stabilize a metalloprotein by –21.6 kcal/mol. Rigorous thermodynamic analyses such as those demonstrated here are critical to current research efforts in metalloprotein design, metal-induced protein folding, and metal-ion trafficking.

Introduction

Metalloprotein engineering involves controlling the delicate interplay between the forces involved in protein folding and the geometric and electronic requirements of the bound metal ion.[1−4] The ever-expanding protein data bank (PDB)[5] provides some insight into how the compendium of natural protein sequences fold to arrange the primary coordination

* To whom correspondence should be addressed. E-mail: brg@chem.columbia.edu.
[†] Columbia University.
[‡] St. John's University.

(1) DeGrado, W. F.; Summa, C. M.; Pavone, V.; Nastri, F.; Lombardi, A. *Annu. Rev. Biochem.* **1999**, *68*, 779−819.
(2) Ghosh, D.; Pecoraro, V. L. *Curr. Opin. Chem. Biol.* **2005**, *9*, 97−103.
(3) Lu, Y. *Curr. Opin. Chem. Biol.* **2005**, *9*, 118−126.
(4) Reedy, C. J.; Gibney, B. R. *Chem. Rev.* **2004**, *101*, 617−650.

10.1021/ic052190q CCC: $33.50
Published on Web 12/04/2006

spheres of bound metals. These ligand constellations contain amino acid side chains, typically in low-energy conformations or rotamers,[6] strategically positioned to provide an appropriate thermodynamic affinity for metal-ion cofactor incorporation and to modulate the metal-ion electronic structure for its biological function.

Of the naturally encoded amino acids used to bind metal ions, cysteine is unique in that its side chain is both redox-active and protonated under standard physiological conditions.[7] In terms of redox activity, the tendency of cysteine to form disulfides, especially in the presence of oxidized metals such as Cu(II),[8] is well-described, as is its ability to be oxidized to thiyl radicals for enzymatic catalysis.[9] In terms of pH, the cysteine thiol is only partially ionized under physiological conditions due to a pK_a value of 8.3.[10] The equilibrium between thiol and thiolate has a significant impact on metal-ion binding because it obscures the inherent value of the metal−ligand formation constant, K_f^{ML}. Consequently, the competition between metals and protons for cysteinate binding at physiological pH weakens the conditional stability constant relative to the inherent K_f^{ML} value of the ligand set.

Despite the redox and pH activity of cysteinate ligands, nature employs thiolate-rich metalloprotein active sites from archaea to higher organisms.[11−15] This observed ubiquity within biological systems emphasizes the operational importance of cysteine-rich metalloprotein active sites, which perform functions ranging from gene expression to enzymatic catalysis.[16−18] Perhaps the most straightforward function of a thiolate-rich metal-ion active site is to utilize the metal−ligand binding thermodynamics to structurally stabilize the protein fold. Zinc finger proteins containing classic structural metal sites with His_xCys_{4-x} ($x = 0-2$) primary coordination spheres have proven critical to our modern understanding of metal-induced protein-folding events.[19,20] In addition to simple structural roles, thiolate-rich iron−sulfur protein active sites provide electron transfer and catalytic function in biology.[21] Understanding the pH and redox dependence of the basic metal−thiolate interaction is crucial to successful metalloprotein engineering efforts and a thorough comprehension of their structural and functional roles in biology.

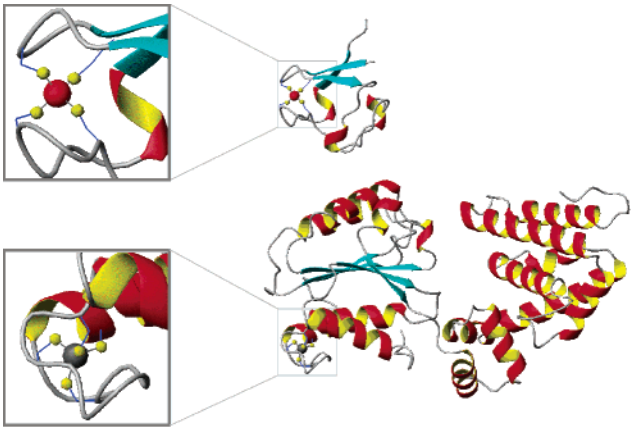


Figure 1. Comparison of global fold and active-site metal-ion geometry of two metalloproteins that utilize tetrathiolate primary coordination spheres to bind mononuclear metal ions. (Top) X-ray crystal structure of *Pf* Rb. The inset shows the Fe, in red, bound by Cys5, Cys8, Cys28, and Cys31.[166] (Bottom) Structure of the γ' subunit of the clamp-loader complex of PolIII, which contains a mononuclear Zn(II), in silver, coordinated by Cys50, Cys59, Cys62, and Cys65.[23] The graphics in Figures 1 and 2 were rendered in *MOLMOL*.[167]

Figure 1 compares the global and active-site structures of two mononuclear proteins with $M(S-Cys)_4$ active sites: *Pyrococcus furiosus* rubredoxin (*Pf* Rb)[22] and the γ' subunit of the clamp-loader complex of *Escherichia coli* DNA polymerase III (PolIII).[23] While the global protein folds are quite distinct, the metal centers share similar metrical parameters within tetrahedral tetracysteinate coordination spheres. In the case of *Pf* Rb, the metal ion is bound between a pair of CxxC sequence motifs remote in the primary structure of the protein but juxtaposed in its tertiary structure. For PolIII, the four cysteinate ligands reside relatively close in the primary structure within a $Cx_8CxxCxxC$ pattern that defines a zinc-binding module between a pair of α helices. The geometric and electronic similarity of these binding sites[22−27] suggests similar metal-ion binding thermodynamics and, thus, a preference of both sites for Zn(II) over Fe(II) based on considerations from the Irving−Williams series[28] or hard−soft acid−base theory.[29] However, recent advances in metallochaperone studies[30−32] have evidenced that biological kinetic control of protein ligands may provide a mech-

(5) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(6) Lovell, S. C.; Word, J. M.; Richardson, J. S.; Richardson, D. C. *Proteins: Struct., Funct., Genet.* **2000**, *40*, 389−408.
(7) Maret, W. *Biochemistry* **2004**, *43*, 3301−3309.
(8) Feldman, S. L.; Hunter, J. S. V.; Zgirski, A.; Chidambaram, M. V.; Frieden, E. *J. Inorg. Biochem.* **1982**, *17*, 51−60.
(9) Licht, S.; Gerfen, G. J.; Stubbe, J. *Science* **1996**, *271*, 477−481.
(10) Gladysheva, T.; Liu, J.; Rosen, B. P. *J. Biol. Chem.* **1996**, *271*, 33256−33260.
(11) Hamer, D. H. *Annu. Rev. Biochem.* **1986**, *55*, 913−951.
(12) Klug, A. *FEBS Lett.* **2005**, *579*, 892−894.
(13) Johnson, D. C.; Dean, D. R.; Smith, A. D.; Johnson, M. K. *Annu. Rev. Biochem.* **2005**, *74*, 247−281.
(14) Auld, D. S. *BioMetals* **2001**, *14*, 271−313.
(15) Vallee, B.; Auld, D. S. *Acc. Chem. Res.* **1983**, *26*, 543−551.
(16) Fomenko, D. E.; Gladyshev, V. N. *Biochemistry* **2003**, *42*, 11214−11225.
(17) Pennella, M. A.; Shokes, J. E.; Cosper, N. J.; Scott, R. A.; Giedroc, D. P. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3713−3718.
(18) Beinert, H.; Holm, R. H.; Münck, E. *Science* **1997**, *277*, 653−659.
(19) Berg, J. M.; Godwin, H. A. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 357−371.
(20) Berg, J. M. *Acc. Chem. Res.* **1995**, *28*, 14−19.
(21) Rao, P. V.; Holm, R. H. *Chem. Rev.* **2004**, *104*, 527−559.
(22) Adman, E. T.; Sieker, L. C.; Jensen, L. H. *J. Biol. Chem.* **1976**, *251*, 3801−3806.
(23) Guenther, B.; Onrust, R.; Sali, A.; O'Donnell, M.; Kuriyan, J. *Cell* **1997**, *91*, 335−345.
(24) Dauter, Z.; Wilson, K. S.; Sieker, L. C.; Moulis, J.-M.; Meyer, J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8836−8840.
(25) Eidsness, M. K.; O'Dell, S. E.; Kurtz, D. M.; Robson, R. L.; Scott, R. A. *Protein Eng.* **1992**, *5*, 367−371.
(26) Jouanneau, Y.; Meyer, C.; Gaillard, J.; Forest, E.; Gagnon, J. *J. Biol. Chem.* **1993**, *268*, 10636−10644.
(27) Arnesano, F.; Banci, L.; Bertini, I.; Ciofi-Baffoni, S. *Eur. J. Inorg. Chem.* **2004**, 1583−1593.
(28) Irving, H.; Williams, R. J. P. *Nature* **1948**, *162*, 746−747.
(29) Pearson, R. G. *J. Am. Chem. Soc.* **1963**, *85*, 3533−3539.
(30) Rosenzweig, A. C. *Chem. Biol.* **2002**, *9*, 673−677.
(31) O'Halloran, T. V.; Culotta, V. C. *J. Biol. Chem.* **2000**, *275*, 25057−25060.
(32) Tottey, S.; Harvie, D. R.; Robinson, N. J. *Acc. Chem. Res.* **2005**, *38*, 775−783.

Table 1. Natural Cysteine-Rich Metal-Ion Binding Sites

metal	protein	protein function	sequence motif	coordination geometry	bite distance (Å)	PDB ID
Fe(II/III)	rubredoxin	electron transfer	$CX_2CX_{29}CX_2C$	tetrahedral	3.81	1BRF
$[2Fe-2S]^{2+/+}$	ferredoxin	electron transfer	$CX_4CX_2CX_{29}C$	tetragonal	3.71	1A70
$[3Fe-4S]^{0/-}$	aconitase	Fe sensor, catalysis	CX_2CX_2C	tetrahedral	6.48	1SJ1
$[4Fe-4S]^{2+/+}$	ferredoxin	electron transfer	$CX_2CX_2CX_3C$	tetrahedral	6.47	1DUR
Cu(I)	CueR	transcriptional activation	CX_7C	digonal	4.45	1Q05
Cu(I)	Hah-1	Cu transport	CX_2C	tetrahedral	4.03	1FEE
Cu(I)	Menkes	Cu transport	CX_2C	trigonal[c]	3.64[a]	1S6U
Zn(II)	DNA polymerase II	catalysis	$CX_2CX_{15}CX_2C$	tetrahedral	3.55	1I50
Zn(II)	DNA polymerase III	catalysis	CX_2CX_2C	tetrahedral	3.64	1A5T
Zn(II)	Ada repair protein	DNA repair	$CX_3CX_{26}CX_2C$	tetrahedral	4.16[a]	1ADN
Zn(II)	metallothionein	metal-ion storage	$CXCX_{13}CX_2C$	tetrahedral	3.96	4MT2
As(III)	ArsR	transcriptional derepression	$CXCX_2C$	trigonal pyramid	3.2[b]	
Hg(II)	MerR	transcriptional activation	CX_8C	trigonal	4.3[b]	

[a] Derived from the solution NMR structure.[168] [b] Estimated based on coordination geometry and EXAFS-derived M−L bond distances.[169] [c] The S−Cu−S angle is indicative of trigonal coordination, but the identity of the exogenous ligand is unknown.

anism with which to circumvent the thermodynamic metal-ion selectivity inherent to tetrahedral tetrathiolate coordination spheres.

The issue of metal-ion selectivity and specificity at cysteine-rich coordination sites[33,34] in protein scaffolds is further complicated by their prevalence in biology and their promiscuity in metal-ion binding. Nature utilizes a CxxC sequence motif both as the functional unit in disulfide redox proteins such as thioredoxin and as a common metal-ion binding unit in metalloproteins.[35] Table 1 lists a selection of proteins that utilize cysteine ligands to bind metal ions which are required for biological function. These sites bind metals ranging from Fe(II) to Hg(II) in digonal, trigonal, tetragonal, and tetrahedral coordination geometries and often include this CxxC motif. This motif is central to the binding of the full variety of iron−sulfur protein active sites including the mononuclear Fe(II), the binuclear [2Fe−2S] cluster, the trinuclear [3Fe−4S] cluster, and both the low- and high-potential tetranuclear [4Fe−4S] clusters.[18] Control of cluster selectivity in iron−sulfur proteins containing CxxC motifs is, in part, determined by the three-dimensional arrangement of their cysteinate ligands. As shown in Figure 2, both the conformation of the cysteine residues, i.e., rotamers,[6] and the resulting ligand bite distance are important parameters in the observed iron−sulfur cofactor selectivity because the various cluster types have distinct coordination geometries. This is also the case in proteins containing metals with specific geometric preferences such as the digonal, or linear, coordination of Cu(I) bound to the CxxC motif in the CueR transcriptional activator.[36]

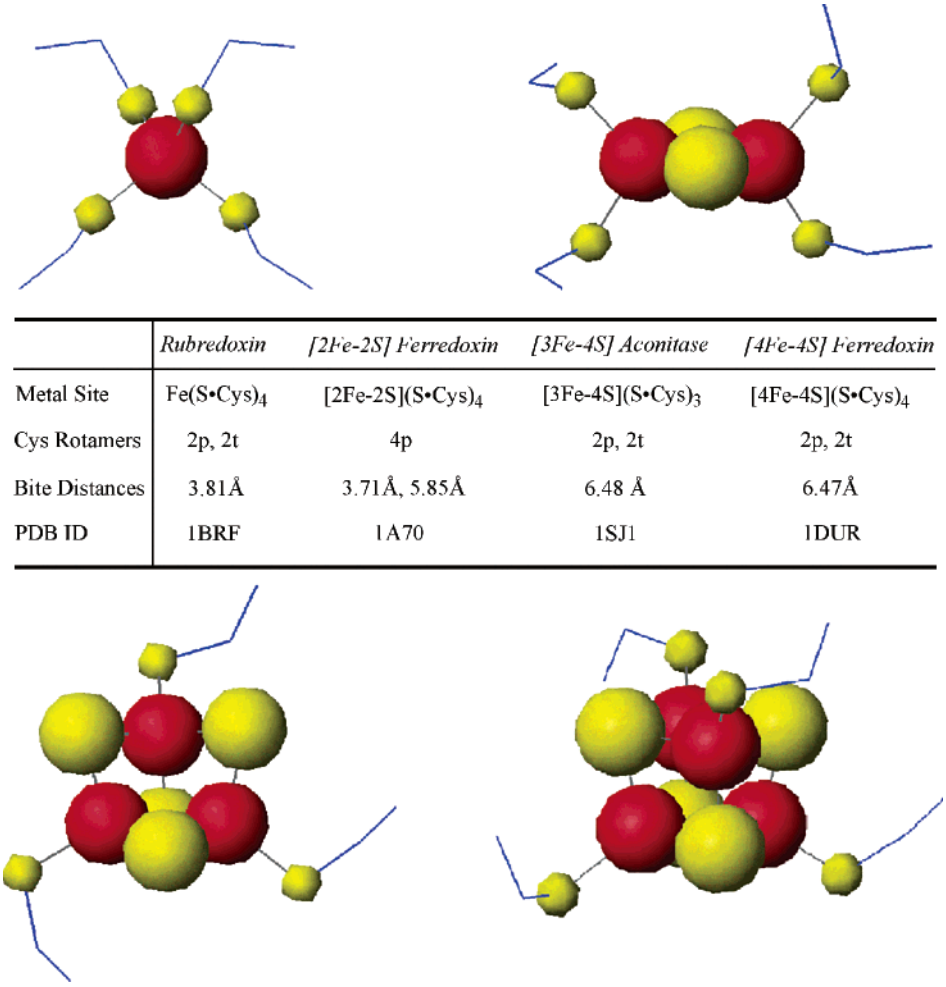
While the ligands and geometry of the primary coordination sphere establish the thermodynamic basis of metal-ion selectivity and specificity, the secondary coordination sphere and beyond are critical in further modulating the metal-ion binding-site affinity and directing the metal's biological function.[37,38] In the case of zinc(II) proteins, the CxxC motif is commonly found in tetrahedral $Zn^{II}(S-Cys)_4$ sites. While many such sites are structural, e.g., zinc ribbons,[39] others serve as sites of reactivity, e.g., the O^6-methylguanine transferase Ada.[40] Secondary-coordination-sphere interactions are critical to controlling cysteinate reactivity at metalloprotein active sites. The lack of an NH···S^{γ} hydrogen bond to Cys38 in the *E. coli* Ada protein activates this ligand to Zn(II) for nucleophilic attack in order to repair methylated DNA.[41]

Peptide-Based Synthetic Analogues of Thiolate-Rich Active Sites

Efforts to understand the geometric and electronic structures, and hence the chemical reactivities, of thiolate-rich active sites in native metalloproteins have expanded from traditional small-molecule bioinorganic chemistry[21,42,43] to the design and synthesis of peptide- and protein-based synthetic analogues.[44,45] The unprecedented success of modeling thiolate-rich protein active sites with simple organic thiolates in the early 1970s led to the expansion into peptide-ligand-based synthetic analogues.[46] Designs in the 1970s and 1980s, which centered on the use of the ubiquitous CxxC motif observed in various metallothioneins and ferredoxins, explored the interaction of cysteine thiolates with metal centers.

In the case of iron−sulfur proteins, minimal peptides have been used as ligands for mononuclear rubredoxin, [2Fe−2S] ferredoxin and [4Fe−4S] ferredoxin active sites. Que et al. synthesized a 9 amino acid peptide, ^{t}Boc-G**C**GG**C**GG**C**G-$CONH_2$, containing 3 cysteines as well as a 12 amino acid

(33) Dudev, T.; Lim, C. *Chem. Rev.* **2003**, *103*, 773−787.
(34) Schymkowitz, J. W. H.; Rousseau, F.; Martins, I. C.; Ferkinghoff-Borg, J.; Stricher, F.; Serrano, L. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 10147−10152.
(35) Buchanan, B. B.; Balmer, Y. *Annu. Rev. Plant Biol.* **2005**, *56*, 187−220.
(36) Changela, A.; Chen, K.; Xue, Y.; Holschen, J.; Outten, C. E.; O'Halloran, T. V.; Mondragón, A. *Science* **2003**, *301*, 1383−1387.
(37) Karlin, S.; Zhu, Z. Y. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 14231−14236.
(38) DiTusa, C. A.; McCall, K. A.; Christensen, T.; Mahapatro, M.; Fierke, C. A.; Toone, E. J. *Biochemistry* **2001**, *40*, 5345−5351.
(39) Qian, X. Q.; Jeon, C. J.; Yoon, H. S.; Agarwal, K.; Weiss, M. A. *Nature* **1993**, *365*, 277−279.
(40) Myers, L. C.; Terranova, M. P.; Ferentz, A. E.; Wagner, G.; Verdine, G. L. *Science* **1993**, *261*, 1164−1167.
(41) He, C.; Hus, J. C.; Sun, L. J.; Zhou, P.; Norman, D. P. G.; Dötsch, V.; Wei, H.; Gross, J. D.; Lane, W. S.; Wagner, G.; Verdine, G. L. *Mol. Cell* **2005**, *20*, 117−129.
(42) Henkel, G.; Krebs, B. *Chem. Rev.* **2004**, *104*, 801−824.
(43) Parkin, G. *Chem. Rev.* **2004**, *104*, 699−767.
(44) Xing, G.; DeRose, V. J. *Curr. Opin. Chem. Biol.* **2001**, *5*, 196.
(45) Kennedy, M. L.; Gibney, B. R. *Curr. Opin. Struct. Biol.* **2001**, *11*, 485.
(46) Herskovitz, T.; Averill, B. A.; Holm, R. H.; Ibers, J. A.; Phillips, W. D.; Weiher, J. F. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 2437−2441.



	Rubredoxin	*[2Fe-2S] Ferredoxin*	*[3Fe-4S] Aconitase*	*[4Fe-4S] Ferredoxin*
Metal Site	$Fe(S•Cys)_4$	$[2Fe\text{-}2S](S•Cys)_4$	$[3Fe\text{-}4S](S•Cys)_3$	$[4Fe\text{-}4S](S•Cys)_4$
Cys Rotamers	2p, 2t	4p	2p, 2t	2p, 2t
Bite Distances	3.81Å	3.71Å, 5.85Å	6.48 Å	6.47Å
PDB ID	1BRF	1A70	1SJ1	1DUR

Figure 2. Comparison of the active-site ligand geometries of tetracysteinate iron−sulfur proteins. An example of each active-site structure type (mono-, bi-, tri-, and tetranuclear) is shown along with the respective ligand bite distances and cysteine rotameric states. The side-chain dihedral, χ_1, values for the plus (p) and trans (t) rotamers of cysteine are +62° and −177°, respectively.[6]

peptide, ^{t}Boc-G**C**GG**C**GG**C**GG**C**G-$CONH_2$, containing 4 cysteines to study $[4Fe-4S]^{2+/+}$ cluster binding and reactivity.[47] In 80% dimethyl sulfoxide in water, both of these peptides bind to a preformed $[4Fe-4S]^{2+}$ cluster as evidenced by UV−visible and NMR spectroscopies. In the case of the 9 amino acid peptide, the fourth coordination site on the peptide-bound [4Fe−4S] cluster is occupied by an exogenous *tert*-butyl thiolate ligand, (S−^{t}Bu), which allows for ligand-exchange reactivity studies. The reduction potentials of these [4Fe−4S] peptide complexes are >300 mV more negative than the ≈−400 mV vs SHE value of natural [4Fe−4S] ferredoxins. This difference may be a simple consequence of comparing electrochemical values between different solvents, including contributions from changes in the liquid-junction potentials.[48]

Aside from [4Fe−4S] cluster incorporation, peptide ligands have also been used to make synthetic analogues of mononuclear rubredoxin and [2Fe−2S] ferredoxin active sites. Anglin and Davison also used the same tetracysteine 12-mer peptide to make the initial peptide-based synthetic analogues of cobalt(II), iron(II), and iron(III) rubredoxin in dimethyl sulfoxide. These complexes were excellent spectroscopic models for rubredoxin in both oxidation states. The data demonstrated that the $Fe^{II}(S-Cys)_4$ chromophore in the metallopeptide possessed a weaker tetrahedral ligand-field splitting, $\Delta_t = 4000-5000\ cm^{-1}$, than that observed in the natural reduced rubredoxin, $\Delta_t = 6250\ cm^{-1}$.[49] Similar observations were made when Ueno et al. used CxxC peptide sequences taken directly from *Clostridium pasteuranium* rubredoxin, Z-**C**PL**C**-OMe or Z-**C**TL**C**-OMe, to coordinate a ferrous ion in dimethylformamide.[50] The identities of the noncoordinating amino acids were found to be important to the electrochemistry of the resulting $\{Fe(S-Cys)_4\}^{-/2-}$ unit. These dicysteine peptides have also been used to generate synthetic analogues of [2Fe−2S] ferredoxins via ligand-exchange reactions to a preformed $[2Fe-2S]^{2+}$ cluster. Similar to the observations of Que et al. with peptide-bound $[4Fe-4S]^{2+/+}$ clusters,[47] the $[2Fe-2S]^{2+/+}$ reduction potentials measured in dimethylformamide were more negative than natural [2Fe−2S] ferredoxins measured in aqueous

(47) Que, L., Jr.; Anglin, J. R.; Bobrik, M. A.; Davison, A.; Holm, R. H. *J. Am. Chem. Soc.* **1974**, *96*, 6042−6048.
(48) Bard, A. J.; Faulkner, L. R. *Electrochemical Methods: Fundamentals and Applications*, 2nd ed.; John Wiley and Sons: New York, 2001; pp 63−74.
(49) Anglin, J. R.; Davison, A. *Inorg. Chem.* **1975**, *14*, 234−237.
(50) Ueno, S.; Ueyama, N.; Nakamura, A.; Tukihara, T. *Inorg. Chem.* **1986**, *25*, 1000−1005.

buffers. In the 1990s, these early iron−sulfur protein synthetic analogues were complemented by the total synthesis of the natural products, a mononuclear rubredoxin and a [4Fe−4S] ferredoxin, by solid-phase peptide synthesis.[51,52] These efforts have advanced to include the incorporation of nonnatural amino acids by native chemical ligation in order to evaluate the role of secondary-coordination-sphere amino acids in rubredoxin and in high-potential iron protein electrochemical function.[53]

Peptide-based synthetic analogues for structural Zn(II) sites have proven valuable in delineating the roles both of the metal ion in protein folding and of the metal−ligand enthalpy in metal-ion selectivity.[54−58] Using sequence alignments of zinc finger transcription factor proteins, Berg and co-workers designed a 26 amino acid consensus zinc finger peptide, CP-1.[55] As was observed in natural zinc fingers, CP-1 is unfolded and no longer exhibits DNA binding activity in the absence of the metal ion. Upon coordination of Zn(II) by the tetrahedral His_2Cys_2 site, the CP-1 sequence folds into its designed $\beta\beta\alpha$ protein structure, which allows for sequence-specific DNA binding functionality. The 5.7 pM conditional dissociation constant, K_d, at pH 7.0 indicates tight thermodynamic binding of Zn(II) to CP-1.[55b] The observed 7.6 kcal/mol selectivity of the protein for Zn(II) over Co(II) at pH 7.0 is postulated to be largely due to the loss of ligand-field stabilization energy (LFSE) as Co(II) goes from an octahedral to a tetrahedral coordination sphere.[59] Furthermore, isothermal titration calorimetry has demonstrated that the binding of Zn(II) is largely enthalpically driven because the entropy of protein folding is compensated for by the release of water from $\{Zn^{II}(H_2O)_6\}^{2+}$ upon binding. The tetracysteinate version of CP-1, CP-1(CCCC), binds Zn(II) with a similar K_d value of 1.1 pM at pH 7.0 despite the change in the primary coordination sphere.[55b] The tight Zn(II) affinity and the metal-induced protein folding of zinc finger proteins were contemporaneously exploited by Imperiali and Berg to generate fluorescent biosensors for Zn(II).[58−60] Imperiali has further improved the dynamic range of these peptide-based sensors, picomolar to nanomolar, by judicious choice of primary-coordination-sphere variations.[61]

De Novo Metalloprotein Design

Over the past 2 decades, the design of folded protein scaffolds from first principles, de novo protein design,[62] has shown considerable success via iterative,[63] combinatorial,[64] and computational design[65] methodologies. Critical to this success was the elucidation of the thermodynamic contribution to global protein fold stabilization by each protein interaction type, e.g., hydrogen bonds, salt bridges, and hydrophobic interactions.[66] These protein-modeling efforts have contributed essential components to our current understanding of protein folding, kinetics and thermodynamics, and protein structure−function relationships.[67] More recently, de novo design has progressed to incorporate metal-ion cofactors in order to access the unique structural and functional properties found in natural metalloproteins.[68] The influence of metal ions in peptide assembly, folding, and stabilization is being delineated in these simplified protein scaffolds.[69] In addition, de novo metalloprotein design seeks to exploit the bound metal ions for ligand-binding,[70] sensing,[71] electron-transfer,[72−74] and catalysis functions.[75,76] Fundamental to the continuing success of these efforts is an elucidation of the metal−ligand thermodynamics relevant to metalloprotein design. The combination of the thermodynamic parameters from metal−ligand interactions with those determined for protein folding is required to provide a complete thermodynamic basis for metalloprotein engineering.

The rational design of stable mononuclear metal-ion sites in de novo designed peptide and protein ligands has shown remarkable success in the past 15 years. Regan and Clarke's initial report[77] of Zn(II) incorporation into a tetrahedral His_2Cys_2 site in α_4, a designed four-α-helix bundle protein,[78]

(51) Christen, R.; Jancic, T.; Zhou, Z. H.; Adams, M. W. W.; Tomich, J. M.; Smith, E. T. *J. Inorg. Biochem.* **1997**, *65*, 53−56.
(52) Smith, E. T.; Feinberg, B. A.; Richards, J. H.; Tomich, J. M. *J. Am. Chem. Soc.* **1991**, *113*, 688−689.
(53) (a) Low, D. W.; Hill, M. G. *J. Am. Chem. Soc.* **1998**, *120*, 11536−11537. (b) Low, D. W.; Hill, M. G. *J. Am. Chem. Soc.* **2000**, *122*, 11039−11040.
(54) Berg, J. M.; Merkle, D. L. *J. Am. Chem. Soc.* **1989**, *111*, 3759−3761.
(55) (a) Krizek, B. A.; Amann, B. T.; Kilfoil, V. J.; Merkle, D. L.; Berg, J. M. *J. Am. Chem. Soc.* **1991**, *113*, 4518−4523. (b) Krizek, B. A.; Merkle, D. L.; Berg, J. M. *Inorg. Chem.* **1993**, *32*, 937−940.
(56) Struthers, M. D.; Cheng, R. P.; Imperiali, B. *Science* **1996**, *271*, 342−345.
(57) Nomura, A.; Sugiura, Y. *Inorg. Chem.* **2004**, *43*, 1708−1713.
(58) Godwin, H. A.; Berg, J. M. *J. Am. Chem. Soc.* **1996**, *118*, 6514−6515.
(59) Blasie, C. A.; Berg, J. M. *Biochemistry* **2002**, *41*, 15068−15073.
(60) Walkup, G. K.; Imperiali, B. *J. Am. Chem. Soc.* **1996**, *118*, 3053−3054.
(61) Walkup, G. K.; Imperiali, B. *J. Am. Chem. Soc.* **1997**, *119*, 3443−3450.
(62) Bryson, J. W.; Betz, S. F.; Lu, H. S.; Suich, D. J.; Zhou, H. X.; O'Neil, K. T.; DeGrado, W. F. *Science* **1995**, *270*, 935−941.
(63) (a) Munson, M.; O'Brien, R. O.; Sturtevant, J. M.; Regan, L. *Protein Sci.* **1994**, *3*, 2015−2022. (b) Gibney, B. R.; Rabanal, F.; Skalicky, J. J.; Wand, A. J.; Dutton, P. L. *J. Am. Chem. Soc.* **1999**, *121*, 4952−4960.
(64) (a) Kamtekar, S.; Schiffer, J. M.; Xiong, H.; Babik, J. M.; Hecht, M. H. *Science* **1993**, *262*, 1680−1685. (b) Davidson, A. R.; Lumb, K. J.; Sauer, R. T. *Nat. Struct. Biol.* **1995**, *2*, 856−864.
(65) (a) Dahiyat, B. I.; Mayo, S. L. *Science* **1997**, *278*, 82−87. (b) Dwyer, M. A.; Looger, L. L.; Hellinga, H. W. *Science* **2004**, *304*, 1967−1971.
(66) (a) Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D. *Science* **2003**, *302*, 1364−1368. (b) Desjarlais, J. R.; Handel *Protein Sci.* **1995**, *4*, 2006−2018.
(67) Hill, R. B.; Raleigh, D. P.; Lombardi, A.; DeGrado, W. F. *Acc. Chem. Res.* **2000**, *33*, 745−754.
(68) Holm, R. H.; Kennepohl, P.; Solomon, E. I. *Chem. Rev.* **1996**, *96*, 2239−2314.
(69) Ghadiri, M. R.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, *114*, 825−831.
(70) Zhuang, J.; Amoroso, J. H.; Kinloch, R.; Dawson, J. H.; Baldwin, M. J.; Gibney, B. R. **2004**, *43*, 8218−8220.
(71) Cerasoli, E.; Sharpe, B. K.; Woolfson, D. N. *J. Am. Chem. Soc.* **2005**, *127*, 15008−15009.
(72) Mutz, M. W.; McLendon, G. L.; Wishart, J. F.; Gaillard, E. R.; Corin, A. F. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 9521−9526.
(73) Sharp, R. E.; Moser, C. C.; Rabanal, F.; Dutton, P. L. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 10465−10470.
(74) Rau, H. K.; DeJonge, N.; Haehnel, W. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11526−11531.
(75) Pinto, A. L.; Hellinga, H. W.; Caradonna, J. P. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 5562−5567.
(76) Moffet, D. A.; Certain, L. K.; Smith, A. J.; Kessel, A. J.; Beckwith, K. A.; Hecht, M. H. *J. Am. Chem. Soc.* **2000**, *122*, 7612−7613.
(77) Regan, L.; Clarke, N. D. *Biochemistry* **1990**, *29*, 10878−10883.
(78) Ho, S. P.; DeGrado W. F. *J. Am. Chem. Soc.* **1987**, *109*, 6751−6758.

demonstrated the feasibility of metalloprotein design from scratch. In the specific case of thiolate-rich metal site design, metalloproteins containing Fe(II), Fe(III), [4Fe−4S], Co(II), Ni(II), Zn(II), As(III), Cd(II), and Hg(II), among others, have been constructed in simple peptide ligands and in folded protein scaffolds.[79−85] The simplest designs are based on linear and cyclic peptides containing CxxC motifs.[79−81] The resulting Fe(II), Co(II), Ni(II), and Zn(II) complexes are excellent synthetic analogues for natural metalloproteins, and their metal-binding preferences are expected to follow the Irving−Williams series. Larger scaffolds designed to bind metal ions in their preferred geometries, such as the tetrahedral Cd(II) site engineered into a two-stranded coiled-coil protein using a pair of CxxC motifs by Kharenko and Ogawa, will likely show similar metal-ion binding preferences.[82] However, protein folding can provide the requisite energy to constrain the binding-site geometry and provide for control of metal-ion selectivity and specificity. Pecoraro and co-workers have pioneered the use of metal−thiolate interactions to drive assembly and folding of de novo designed helical bundle scaffolds.[83] Three identical peptides, each containing a single cysteine residue, assemble, fold, and bind a Hg(II), As(III), or Cd(II) ion within the protein hydrophobic core in an aqueous solution.[84,85] Detailed kinetic and thermodynamic studies on Hg(II) binding demonstrate the role of protein assembly and folding in controlling metal-ion coordination geometry. The digonal $Hg^{II}(S-Cys)_2$ coordination preferred by the metal ion can be driven by solution pH to a trigonal $Hg^{II}(S-Cys)_3$ coordination unit preferred by the protein fold.[84]

[4Fe−4S] Protein Design

Dutton and co-workers have dubbed minimal de novo designed metalloproteins as *maquettes*, an architectural term denoting a smaller model of a larger work.[86−88] Their initial protein maquette, $[H10H24]_2$, is a synthetic four-α-helix bundle scaffold that binds four hemes via linear bis(histidine) coordination within the helical regions. Extension of the maquette concept to thiolate-rich metalloprotein design provided the initial ferredoxin and ferredoxin−heme maquettes.[89] The prototype ferredoxin maquette, FdM, is a 16 amino acid peptide designed to employ four cysteines to coordinate a tetranuclear iron−sulfur cofactor, as shown for the related **IGA** peptide in Figure 3. This initial design uses the consensus bacterial ferredoxin [4Fe−4S] binding motif of −CxxCxxC− to provide three of the four cysteine ligands. The −CIACGAC− consensus motif from the gram-positive anaerobe *Peptostreptococcus asaccharolyticus* ferredoxin I[22] was selected for the FdM. As such, this design deviates from the purest definition[1] of de novo protein design in that it recapitulates a natural sequence. The fourth cysteine was placed four amino acid positions N-terminal from the consensus motif CEGG−CIACGAC− in accordance with in silico modeling. The resulting sequence, [CEGG−CIAC]-GAC−, also contains the [CxxxCxxC] cysteine spacing found in *S*-adenosylmethionine (AdoMet)-dependent iron−sulfur enzymes such as pyruvate formate lyase activating enzyme[90] and biotin synthase.[91] In addition to these metal-ion binding residues, a single tryptophan on a glycine tail was included at the C terminus as a spectroscopic tag, and lysine and leucine were placed on the N terminus to facilitate incorporation into a helix−loop−helix ferredoxin−heme maquette.[89] The resulting FdM sequence is H_2N-KLCEGG−C**IA**CG**A**C−GGW-$CONH_2$ and will be referred to hereafter as **IAA** for the amino acids in positions 8, 9, and 12, respectively.

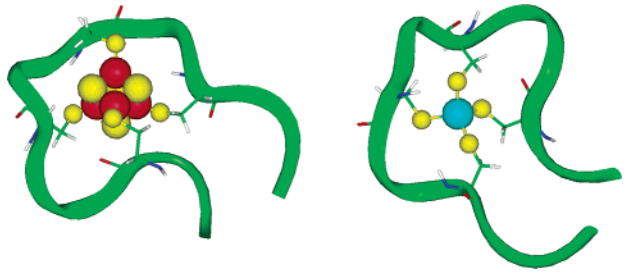


Figure 3. Ribbon diagrams of the $[4Fe-4S]^{2+}$-**IGA** (left) and Co^{II}-**IGA** (right) molecular models showing the metal centers and the cysteinate ligands. The [4Fe−4S] cluster in red/yellow and Co(II) in light blue are bound by a tetrahedral array of cysteine thiolates with bite distances of 6.8 Å and 3.8 Å, respectively. The models indicate that there is no steric hindrance that prevents the binding of a mononuclear metal center despite the smaller bite distance required.

Under strictly anaerobic conditions and in the presence of β-mercaptoethanol (β-ME) as a sacrificial reductant, a single [4Fe−4S] cluster can be incorporated into the **IAA** ligand in aqueous buffers using the self-assembly reaction in eq 1.

$$\textbf{IAA}\text{-}4H^+ + 4FeCl_3 + 4Na_2S + 2e^- \rightleftharpoons [4Fe-4S]^{2+}\text{-}\textbf{IAA} + 8NaCl + 4HCl$$

As is evident from the reaction in eq 1, [4Fe−4S] binding is expected to be highly pH dependent because the **IAA**-$4H^+$ ligand loses four protons upon metal-ion coordination.[92] The assembly of the cluster is independent of the order of addition of $FeCl_3$ and Na_2S to **IAA** and presumably proceeds

(79) Kennedy, M. L.; Gibney, B. R. *J. Am. Chem. Soc.* **2002**, *124*, 6826−6827.
(80) Petros, A. K.; Shaner, S. E.; Costello, A. L.; Tierney, D. L.; Gibney, B. R. *Inorg. Chem.* **2004**, *43*, 4793−4795.
(81) Nivorozhkin, A. L.; Segal, B. M.; Musgrave, K. B.; Kates, S. A.; Hedman, B.; Hodgson, K. O.; Holm, R. H. *Inorg. Chem.* **2000**, *39*, 2306−2313.
(82) Kharenko, O. A.; Ogawa, M. Y. *J. Inorg. Biochem.* **2004**, *98*, 1971−1974.
(83) Dieckmann, G. R.; McRorie, D. K.; Tierney, D. L.; Utschig, L. M.; Singer, C. P.; O'Halloran, T. V.; Penner-Hahn, J. E.; DeGrado, W. F.; Pecoraro, V. L. *J. Am. Chem. Soc.* **1997**, *119*, 6195−6196.
(84) Farrer, B. T.; Harris, N. P.; Balchus, K. E.; Pecoraro, V. L. *Biochemistry* **2001**, *40*, 14696−14705.
(85) Ghosh, D.; Lee, K. H.; Demeler, B.; Pecoraro, V. L. *Biochemistry* **2005**, *44*, 10732−10740.
(86) Robertson, D. E.; Farid, R. S.; Moser, C. C.; Urbauer, J. L.; Mulholland, S. E.; Pidikiti, R.; Lear, J. D.; Wand, A. J.; DeGrado, W. F.; Dutton, P. L. *Nature* **1994**, *368*, 425−431.
(87) Choma, C. T.; Lear, J. D.; Nelson, M. J.; Dutton, P. L.; Robertson, D. E.; DeGrado, W. F. *J. Am. Chem. Soc.* **1994**, *116*, 856−865.
(88) Reedy, C. J.; Kennedy, M. L.; Gibney, B. R. *Chem. Commun.* **2003**, 570−571.
(89) Gibney, B. R.; Mulholland, S. E.; Rabanal, F.; Dutton, P. L. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 15041−15046.
(90) Cheek, J.; Broderick, J. B. *J. Biol. Inorg. Chem.* **2001**, *6*, 209−226.
(91) Berkovitch, F.; Nicolet, Y.; Wan, J. T.; Jarrett, J. T.; Drennan, C. L. *Science* **2004**, *303*, 76−79.

via the formation of $\{[4Fe-4S]^{2+}(\beta\text{-ME})_4\}^{2-}$ followed by thiolate ligand exchange with the **IAA** scaffold. Restricting the amount of iron and sulfide in an attempt to generate a [2Fe−2S] cluster results in 50% yield of the [4Fe−4S]-**IAA** complex, consistent with the findings of Rabinowitz et al.[93] in the all-or-none reconstitution reactions of natural ferredoxins. In addition to the lack of observable binuclear [2Fe−2S] incorporation, mononuclear rubredoxin and trinuclear [3Fe−4S] cluster binding are not observed under the reconstitution conditions employed.

[4Fe−4S]-**IAA** and related complexes have served as effective synthetic analogues for bacterial ferredoxins under buffered aqueous conditions. In the $[4Fe-4S]^{2+}$ state, the UV−visible and circular dichroism spectra are indicative of a peptide-bound $[4Fe-4S]^{2+}$ cluster. Reduction to the $[4Fe-4S]^{+}$ state leads to the observation of an electron paramagnetic resonance (EPR) signal whose *g* values and temperature dependence are definitive of a tetranuclear iron−sulfur cluster.[94] The midpoint reduction potential of the $[4Fe-4S]^{2+/+}$-**IAA** complex is directly comparable to that of *P. asaccharolyticus* ferredoxin I because both were measured in aqueous buffers. The reduction potential of $[4Fe-4S]^{2+/+}$-**IAA** measured by EPR spectroelectrochemistry is −350 mV vs SHE at pH 8, which is only slightly more positive than the −430 mV value observed at pH 7.65 for *P. asaccharolyticus* ferredoxin I.[95]

Contemporaneous to the advent of ferredoxin maquettes, two other iron−sulfur protein designs appeared in the literature that do not utilize the clostridial ferredoxin CxxCxxC motif. In the first example, Scott and Biggins[96] used an approach similar to that of the ferredoxin maquette to design a [4Fe−4S] binding site between a pair of loops in α_4, the de novo designed four-α-helix bundle that had been previously used to incorporate Zn(II).[77] Their design, α_4-FeS, replaced two interhelical loops in α_4 with the [4Fe−4S] binding sequence of F_x from the photosystem I reaction center, −P**C**DGPGRGGT**C**−. The resulting α_4-FeS protein bound a $[4Fe-4S]^{2+/+}$ cluster between the two grafted loop regions, as demonstrated by UV−vis and EPR spectroscopies with a midpoint reduction potential of −422 mV at pH 8.3. The spectroscopic and electrochemical properties of α_4-FeS and $[4Fe-4S]^{2+/+}$-**IAA** are quite similar. In a second example, Coldren et al. employed a computational approach to the design of a [4Fe−4S] binding site in the hydrophobic core of a natural protein.[97] The program *DEZYMER*[98] was used to search the structure of thioredoxin, Trx, for a set of mutations that would generate a cavity isosteric with a [4Fe−4S] unit surrounded by four cysteine ligands in low-energy rotamers. While the resulting sequence of cysteine ligands does not correspond to any known consensus [4Fe−4S] binding motif, a [4Fe−4S] cluster could be placed into the Trx-Fe_4S_4 protein by a series of ligand displacement reactions from the preformed $\{[4Fe-4S](\beta\text{-ME})_4\}^{2-}$ cluster. The resulting $[4Fe-4S](S-Cys)_4$ cluster in Trx-Fe_4S_4 showed EPR and electrochemical characteristics of high-potential iron−sulfur proteins (HiPIPs). The burial of this cluster into a low-dielectric hydrophobic core is likely to protect it from hydrolysis and stabilize the oxidized HiPIP state because it carries less overall formal charge, e.g., $\{[4Fe-4S]^{3+}\text{-}(S-Cys)_4\}^{-}$. This situation mirrors the use of sterically crowded thiolates to generate stable HiPIP-type $[4Fe-4S]^{3+/2+}$ cluster complexes in small-molecule bioinorganic chemistry.[99]

Loop-Based Metalloprotein Design

The concept of grafting metal-ion binding loops from one protein to another is not limited to designed protein scaffolds. Lu and co-workers have elegantly utilized loop-directed mutagenesis to install novel metal-ion binding sites into natural protein scaffolds.[100] Most notably, splicing the Cu_A binding loop from cytochrome *c* oxidase into the blue copper protein azurin is sufficient to generate a binuclear Cu_A-like site. The conversion of a mononuclear site into a binuclear center in an otherwise invariant scaffold has provided for a direct comparison of their structures and functions. The rate of electron transfer to the binuclear site is enhanced 3-fold relative to the mononuclear site because of a decrease in the reorganization energy,[101] and the reduction potential of the binuclear site is less sensitive to replacement of an axial methionine residue.[102] Using a similar design concept, Franklin and co-workers have installed a Ca(II) binding sequence, an EF-hand motif, between a pair of designed helical sequences to generate a series of lanthanide ion based artificial nucleases.[103]

The reader should not be left with the impression from these examples that the design of simple metalloproteins by excision of metal-ion binding loops from their natural protein contexts is always entirely successful. This is especially true in cases where protein folding and hydrophobic burial exert control over metal-ion binding preferences, i.e., entatic[104] or rack state[105] active sites. For example, attempts to design a synthetic blue copper protein using a variety of design concepts have failed to produce the S → Cu(II) ligand-to-

(92) The peptides in the tetrathiolate state is designated **IGA** and **IGA**, while the peptides with four protonated cysteine thiols are referred to as **IAA**-$4H^+$ and **IGA**-$4H^+$ throughout the manuscript. For clarity, the formal charges on the peptide complexes are not shown.

(93) Malkin, R.; Rabinowitz, J. C. *Biochem. Biophys. Res. Commun.* **1966**, *23*, 822−827.
(94) Cammack, R. Cyanobacteria. In *Methods in Enzymology*; Packer, L., Glazer, A. N., Eds.; Elsevier: London, U.K., 1988; Vol. 167, pp 427−436.
(95) Stombaugh, N. A.; Sundquist, J. E.; Burris, R. H.; Orme-Johnson, W. H. *Biochemistry* **1976**, *15*, 2633−2641.
(96) Scott, M. P.; Biggins, J. *Protein Sci.* **1997**, *6*, 340−346.
(97) Coldren, C. D.; Hellinga, H. W.; Caradonna, J. P. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 6635−6640.
(98) Hellinga, H. W.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 763−785.
(99) O'Sullivan, T.; Millar, M. M. *J. Am. Chem. Soc.* **1985**, *107*, 4096−4097.
(100) Hay, M.; Richards, J. H.; Lu, Y. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 461−464.
(101) Farver, O.; Lu, Y.; Ang, M. C.; Pecht, I. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 899−902.
(102) Berry, S. M.; Ralle, M.; Low, D. W.; Blackburn, N. J.; Lu, Y. *J. Am. Chem. Soc.* **2003**, *125*, 8760−8768.
(103) Kovacic, R. T.; Welch, J. T.; Franklin, S. J. *J. Am. Chem. Soc.* **2003**, *125*, 6656−6662.
(104) Vallee, B. L.; Williams, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **1968**, *59*, 498−505.
(105) Gray, H. B.; Malmström, B. G.; Williams, R. J. P. *J. Biol. Inorg. Chem.* **2000**, *5*, 551−559.

metal charge-transfer (LMCT) band characteristic of the natural proteins azurin and plastocyanin.[106−109] Using a design concept similar to that of the ferredoxin maquettes, Dougherty et al. utilized the plastocyanin metal-ion binding sequence, Tyr**Cys**SerPro**His**GlnGlyAlaGly**Met**ValGlyLys, in a blue copper protein design, BCP-A.[109] In addition to the motif containing the cysteine, histidine, and methionine ligands, the fourth ligand, a histidine, was placed four amino acids C-terminal to the methionine residue and a tryptophan was added on a N-terminal WGGGS sequence as a spectroscopic marker. In pH 9.8 buffers containing NH_3, BCP-A bound Cu(II) in a distorted square-planar geometry using the cysteine and histidine ligands with weak axial solvation. At lower pH values, cysteine oxidation was problematic, and at higher pH values, $\{Cu(NH_3)_4\}^{2+}$ formation was favored, demonstrating the difficulty of aqueous copper coordination chemistry with cysteine thiolates. A second design that placed a β turn between the potential ligands, BCP-B or NH_2-Trp(Gly)$_3$**Cys**Gly**His**GlyValProSer**His**Gly**Met**Gly-$CONH_2$, showed enhanced water stability for the tetragonal Cu(II) complex with axial solvation. The results of this study emphasized the role of natural blue copper protein scaffolds in protecting the Cu(II) site from solvent exposure and cysteine oxidation. Other attempts to generate blue copper proteins in protein scaffolds using rational, combinatorial, and computational design methods have encountered similar issues of cysteine oxidation and the square-planar preferences of the Cu(II) ion.[106−108]

Multicofactor Metalloprotein Design

In cases where loop excision and grafting are successful, they provide an entrée into the modular design of complex metalloproteins, i.e., those containing two or more distinct metal cofactors. In natural proteins, the ability to combine redox cofactors provides for the design of not only the electron-transfer chains but also enzyme active sites capable of multielectron catalysis, e.g., cytochrome *c* oxidase, photosystems I and II, and nitrogenase.[110−113] In the case of the ferredoxin maquette, the peptide termini of the **IAA** sequence were specifically designed to couple to the helices of the heme protein maquette $[H10H24]_2$ to demonstrate the feasibility of modular metalloprotein design.[89] The C-terminal CGGW sequence and the N-terminal KLCE sequence of **IAA** were based on the CGGG loop and the N-terminal ELWK sequence of $[H10H24]_2$, respectively. For the N-terminal sequence, molecular models suggested that a Leu → Cys alteration would best present the Cys−S$^{\gamma}$ for [4Fe−4S] coordination and that the Trp → Leu change was needed to allow for [4Fe−4S] binding. The resulting 67 amino acid sequence was designed to contain two digonal His_2 heme binding sites within the helices and a single tetrahedral Cys_4 [4Fe−4S] binding site in the loop of the helix−loop−helix monomer. In solution, the ferredoxin−heme protein maquette folds and associates as a dimer, or a four-α-helix bundle, prior to binding a pair of [4Fe−4S] clusters and four hemes. The ability to combine hemes with iron−sulfur clusters in the same designed protein provides for the future design of maquettes for succinate dehydrogenase[114] and sulfite reductase.[115] Thus, the thermodynamic preferences of the metal-ion cofactors for distinct coordination sites can be used to drive the high-fidelity self-assembly of complex metalloproteins.

The modular metalloprotein design concept has been extended by Laplaza and Holm to provide for the rational design of coupled metalloprotein active sites.[116] A cysteine thiolate bridged $Ni^{II}[4Fe-4S]^{2+}$ center has been designed into a helix−loop−helix scaffold similar to the ferredoxin−heme protein maquette. In addition to a −CIACGAC− sequence in the loop region, the helical regions contained either a His_3Cys_1 or a His_2Cys_2 site designed to bind Ni(II) and provide the cysteine thiolate bridge to a $[4Fe-4S]^{2+}$ cluster. Detailed extended X-ray absorption fine structure (EXAFS) analysis of the products demonstrated the presence of the desired $\{Ni^{II}$-μ_2-Cys-$[4Fe-4S]^{2+}\}$ units that reproduce the connectivity of early models of the A cluster in carbon monoxide dehydrogenase.[117] The design of these sophisticated protein-based synthetic analogues relies on the inherent coordination chemistry of the metal-ion cofactors.

Protein Engineering Requirements for Ferredoxin Maquette Assembly and Stability

Aside from their applicability as modular components of multicofactor designed proteins, the ferredoxin maquettes have proven their utility as synthetic analogues in delineating the role of the peptide sequence in modulating the stability of the bound [4Fe−4S] cluster.[118,119] In terms of the primary coordination sphere, the inherent thermodynamic preference of the $[4Fe-4S]^{2+/+}$ cluster for the thiolate ligands over carboxylates and imidazoles seen in small-molecule synthetic analogues is also observed in a series of ferredoxin maquettes containing potential carboxylate and histidine ligands.[118] Where ferredoxin maquettes have revealed novel insight into natural ferredoxin engineering is in elucidating the role of the nonligating amino acids in [4Fe−4S] stabilization and electrochemical function.[119]

The nonligating amino acids of **IAA** are just as critical to stabilizing the bound [4Fe−4S] cluster as are the amino acids

(106) Hellinga, H. W. *J. Am. Chem. Soc.* **1998**, *120*, 10055−10066.
(107) Schnepf, R.; Horth, P.; Bill, E.; Wieghardt, K.; Hildebrandt, P.; Haehnel, W. *J. Am. Chem. Soc.* **2001**, *123*, 2186−2195.
(108) Schnepf, R.; Haehnel, W.; Wieghardt, K.; Hildebrandt, P. *J. Am. Chem. Soc.* **2004**, *126*, 14389−14399.
(109) Daugherty, R. G.; Wasowicz, T.; Gibney, B. R.; DeRose, V. J. *Inorg. Chem.* **2002**, *41*, 2623−2632.
(110) Brettel, K. *Biochim. Biophys. Acta* **1997**, *1318*, 322−373.
(111) Kim, J.; Rees, D. C. *Biochemistry* **1994**, *33*, 389−397.
(112) Capaldi, R. A. *Annu. Rev. Biochem.* **1990**, *59*, 569−596.
(113) Barber, J. *Curr. Opin. Struct. Biol.* **2002**, *12*, 523−530.
(114) Yankovskaya, V.; Horsefield, R.; Tornroth, S.; Luna-Chavez, C.; Miyoshi, H.; Leger, C.; Byrne, B.; Cecchini, G.; Iwata, S. *Science* **2003**, *299*, 700−704.
(115) Crane, B. R.; Siegel, L. M.; Getzoff, E. D. *Science* **1995**, *270*, 59−67.
(116) Laplaza, C. E.; Holm, R. H. *J. Am. Chem. Soc.* **2001**, *123*, 10255−10264.
(117) Musgrave, K. B.; Laplaza, C. E.; Holm, R. H.; Hedman, B.; Hodgson, K. O. *J. Am. Chem. Soc.* **2002**, *124*, 3083−3092.
(118) Mulholland, S. E.; Gibney, B. R.; Rabanal, F.; Dutton, P. L. *J. Am. Chem. Soc.* **1998**, *120*, 10296−10302.
(119) Mulholland, S. E.; Gibney, B. R.; Rabanal, F.; Dutton, P. L. *Biochemistry* **1999**, *38*, 10442−10448.

in the primary coordination sphere. This is clearly shown when all of the nonligating amino acids in the **IAA** maquette are replaced with glycines, FdM-Gly, which produces a ligand similar to the 12 amino acid peptide of Que et al.[47] While both **IAA** and FdM-Gly bind a single $[4Fe-4S]^{2+}$ cluster in high yield, the reduced-state EPR spectrum of [4Fe−4S]-FdM-Gly was found to be 10-fold less intense than that of the [4Fe−4S]-**IAA** prototype. These EPR data indicate significant destabilization of the $[4Fe-4S]^{+}$ state in FdM-Gly. Making a single glycine to isoleucine change in FdM-Gly improved the EPR intensity of the reduced state cluster 5-fold, demonstrating that isoleucine was critical for reduced cluster stability. The observation of the importance of isoleucine to cluster stability is reflected in a bioinformatics analysis[120] of iron−sulfur protein sequences that contain the CxxCxxC motif. Analysis of 510 ferredoxin sequences showed that the second position of the motif (position 8 of the ferredoxin maquette) is dominated by the β-branched amino acids isoleucine (246 occurrences), valine (94), and threonine (64).[119] Because the hydrophobic isoleucine residue in *P. asaccharoyliticus* ferredoxin I lies over one face of the cluster and provides a $NH\cdots S^{\gamma}$ hydrogen bond, perhaps its role in the ferredoxin maquette is to protect the reduced cluster from hydrolysis. In addition to identifying the role of the isoleucine residue in the classic [4Fe−4S] motif, the data also demonstrate that glycine is the most common amino acid in position 5 (11 of the maquette). The prevalence of glycine at this position is likely due to its ability to assume the non-Ramachandran torsion angles observed in the structures of natural ferredoxins.

While the prototype **IAA** ferredoxin maquette utilized the −C**IA**CG**A**C− sequence from *P. asaccharolyticus*, the bioinformatics data indicate that −C**IG**CG**A**C− is more commonly observed in bacterial ferredoxins. Kennedy and Gibney have shown that a ferredoxin maquette based on this sequence, **IGA** or H_2N-KLCEGG-C**IG**CG**A**C-GGW$-CONH_2$, binds a $[4Fe-4S]^{2+/+}$ cluster with spectroscopic properties virtually identical with those of the [4Fe−4S]-**IAA** prototype.[79] The most noteworthy difference between the two is a 45 mV, or 1.1 kcal/mol, elevation of the midpoint reduction potential of [4Fe−4S]-**IGA** relative to [4Fe−4S]-**IAA** at pH 7.5.

The reduction potential of [4Fe−4S]-**IGA** is pH-dependent. A plot of the midpoint reduction potential of [4Fe−4S]-**IGA** versus the solution pH, a Pourbaix diagram, demonstrates the binding of a single proton upon $[4Fe-4S]^{2+}$-**IGA** reduction. The data indicated a reduced-state proton dissociation constant, pK_a^{red}, value of 9.3 ± 0.1. While the oxidized-state proton dissociation constant, pK_a^{ox}, value of the peptide-bound cluster could not be evaluated from the electrochemical data because of cluster decomposition under acidic conditions, a limiting pK_a^{ox} value of 6.35 ± 0.05 was determined by direct competition of $[4Fe-4S]^{2+}$-**IGA** with protons. In addition, the shift of this pK_a^{red} value to 8.3 ± 0.1 observed upon selenium substitution into the cluster, $[4Fe-4Se]^{+}$-**IGA**, indicates that the thermodynamics of protonation are sensitive to the cluster type. These data further suggest that protonation occurs near the cluster, perhaps on a cysteinate or μ_3-sulfido ligand. This observation of proton-coupled electron transfer[121] in the ferredoxin maquettes, a feature uncommon to [4Fe−4S] proteins but sometimes observed in [3Fe−4S] proteins,[122] may be one consequence of excising the cofactor from its native context. This suggests that one role of the extensive hydrogen-bond network in natural [4Fe−4S] ferredoxins is to act as a pH buffer for the cluster. Slight changes in hydrogen bonding could act as the charge-compensation mechanism for natural [4Fe−4S] proteins during oxidation/reduction.[123] In the absence of this natural buffering capacity, the $[4Fe-4S]^{2+}$ cluster in the maquette acquires a proton from solution upon reduction. Thus, future [4Fe−4S] protein designs may need to pay closer attention to the design of hydrogen-bonding patterns to achieve pH-independent redox activity.

Metal-Ion Selectivity and Specificity in Thiolate-Rich Designed Proteins

Elucidation of the metal-binding thermodynamics using coordination chemistry methods[124] provides the fundamental energetics relevant to the field of not only metalloprotein design[4] but also metal-induced protein folding[125] and metal-ion storage, uptake, and delivery.[32] Having employed the ferredoxin maquettes to reveal the essential engineering requirements of [4Fe−4S] proteins, herein we use these synthetic analogues to investigate the thermodynamics of thiolate-rich metalloprotein active-site metal-ion specificity and selectivity.[126] The random-coil structure of the ferredoxin maquette presents a model system in which to study the inherent thermodynamic preferences of a Cys_4 metal-ion binding site in the absence of protein-folding contributions. These basic thermodynamic data are needed in the field of metalloprotein design to evaluate the energetics between metal-ion binding and protein folding. Without these data, predictions of the affinities of tetrahedral tetrathiolate peptide ligands for two metal cofactors cannot be made with any accuracy, which has hindered the development of computational methods for predicting the affinity of metal-ion binding sites.[34]

Herein, we investigate the binding of Fe(II), Co(II), Zn(II), and $[4Fe-4S]^{2+}$ to a simple tetradentate tetrathiolate peptide ligand, **IGA** shown in Figure 3, to provide the thermodynamics of metal−ligand binding requisite for future metalloprotein design efforts. Conditional stability constants for Fe(II), Co(II), and Zn(II) are determined at various pH values

(120) The Expert Protein Analysis System Proteomics Server, http://ca.expasy.org/.
(121) Cukier, R. I.; Nocera, D. G. *Annu. Rev. Phys. Chem.* **1998**, *49*, 337−369.
(122) Camba, R.; Jung, Y. S.; Hunsicker-Wang, L. M.; Burgess, B. K.; Stout, C. D.; Hirst, J.; Armstrong, F. A. *Biochemistry* **2003**, *42*, 10589−10599.
(123) Adman, E.; Watenpaugh, K. D.; Jensen, L. H. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 4854−4858.
(124) Magyar, J. S.; Godwin, H. A. *Anal. Biochem.* **2003**, *320*, 39−54.
(125) Struthers, M. D.; Cheng, R. P.; Imperiali, B. *J. Am. Chem. Soc.* **1996**, *118*, 3073−3081.
(126) Kennedy, M. L.; Petros, A. K.; Gibney, B. R. *J. Inorg. Biochem.* **2004**, *98*, 727−732.

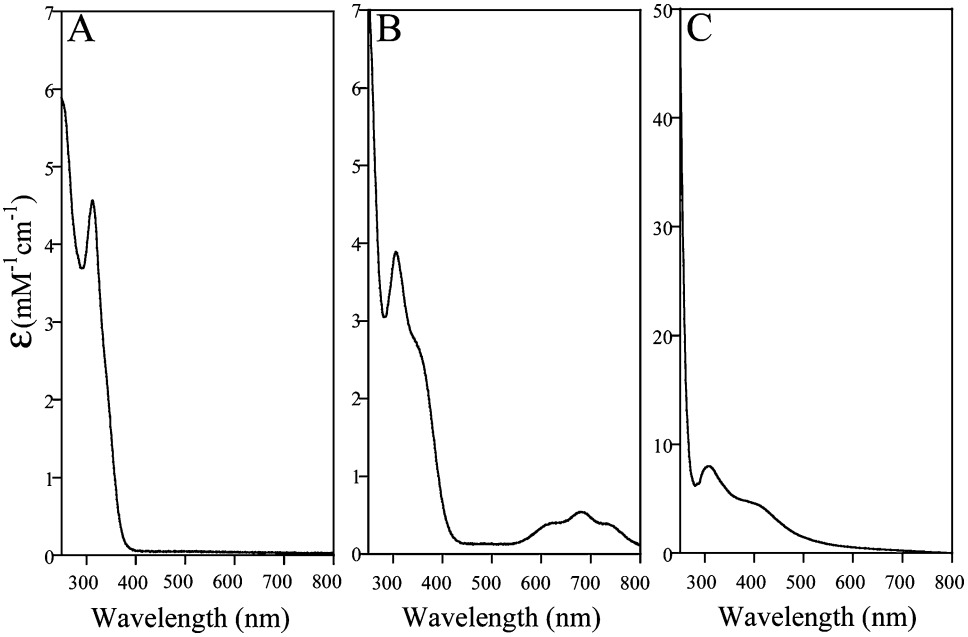


Figure 4. UV−visible spectral comparison of the (A) Fe^{II}-**IGA**, (B) Co^{II}-**IGA**, and (C) $[4Fe-4S]^{2+}$-**IGA** complexes. All complexes were made from a freshly prepared 62 μM stock solution of **IGA** in a 20 mM HEPES buffer containing 100 mM KCl at pH 7.5.

to illustrate the proton sensitivity of metal−peptide complex formation. A rigorous description of the influence of protons on the thermodynamic stabilities of these metallopeptides is presented in order to fully elucidate their aqueous coordination chemistry. The acid dissociation constant values of the **IGA** cysteines bound to each metal ion, pK_a^{eff} values, are determined by direct proton competition experiments and reveal the thermodynamic stabilities of each metal−ligand complex as a function of the pH. Comparison of the inherent formation constants, K_f^{ML} or β_{110} values, of the mononuclear Fe^{II}-, Co^{II}-, and Zn^{II}-**IGA** complexes correlates well with qualitative expectations based on the Irving−Williams series, Fe(II) < Co(II) < Zn(II).[28] Remarkably, the data illustrate that **IGA** binds Zn(II) with a femtomolar affinity that rivals even the tightest Zn(II) binding sites found in folded natural proteins.[127] In addition, the pH stability of the $[4Fe-4S]^{2+}$-**IGA** complex allows us to tentatively suggest a location for the $[4Fe-4S]^{2+}$ complex ion in the Irving−Williams series and provide a thermodynamic rationale for the preferred formation of $[4Fe-4S]^{2+}$-**IGA** over Fe^{II}-**IGA** under cluster self-assembly conditions. These data report the binding preferences of the Cys_4 ligand in the absence of significant protein-folding effects. Comparison of these data to those from natural protein scaffolds is used to reveal the influence of metal-ion binding on the thermodynamics of protein folding.[128]

Thermodynamic Analysis of the Fe(II) Affinity of IGA

Figure 4A shows the UV−visible spectrum of Fe^{II}-**IGA** formed by the addition of 1.0 equiv of ferrous ammonium sulfate to 50 μM **IGA** at pH 7.5 (20 mM HEPES buffer, 100 mM KCl). The spectrum exhibits a LMCT band at 314 nm (ϵ = 4600 M^{-1} cm^{-1}) with a shoulder at 340 nm (ϵ = 2900 M^{-1} cm^{-1}) that is characteristic of mononuclear ferrous tetrathiolate binding sites. The spectrum is similar to those observed in small-molecule inorganic models such as $\{Fe^{II}(S-Ph)_4\}^{2-}$ and $\{Fe^{II}(S-Et)_4\}^{2-}$,[129,130] other designed $Fe^{II}(S-Cys)_4$ peptides,[131−134] and natural proteins, including the rubredoxin from *Desulfovibrio desulfuricans*[135] and ferrous iron bound to metallothionen.[136] Thus, the Fe^{II}-**IGA** complex serves as a very good spectroscopic model for reduced rubredoxin.

To fully describe the thermodyamic stability of Fe^{II}-**IGA**, three types of experiments were conducted. First, conditional dissociation constant values for the Fe^{II}-**IGA** complex were measured by direct metal-ion titration over the pH range of 6.8−7.4 using UV−visible spectroscopy. As expected, weaker K_d values were measured as the concentration of protons was increased. Second, the competition constant for the displacement of Fe(II) in Fe^{II}-**IGA** by Co(II), $K_{comp}^{Fe/Co}$, was determined at pH 7.5. Third, the stability of the Fe^{II}-**IGA** complex with respect to protonation of the cysteine thiolates was determined by direct acid titration. Collectively examining the results from these three methods provides the K_d values of the Fe^{II}-**IGA** complex over a wide range of pH

(127) (a) Hitomi, Y.; Outten, C. E.; O'Halloran, T. V. *J. Am. Chem. Soc.* **2001**, *123*, 8614−8615. (b) Outten, C. E.; O'Halloran, T. V. *Science* **2001**, *292*, 2488−2492.
(128) Maret, W.; Vallee, B. L. *Metallobiochemistry, Part C* **1993**, *226*, 52−71.
(129) Coucouvanis, D.; Swenson, D.; Baenziger, N. C.; Murphy, C.; Holah, D. G.; Sfarnas, N.; Simopoulos, A.; Kostikas, A. *J. Am. Chem. Soc.* **1981**, *103*, 3350−3362.
(130) Koch, S. A.; Maelia, L. E.; Millar, M. *J. Am. Chem. Soc.* **1983**, *105*, 5944−5945.
(131) Farinas, E.; Regan, L. *Protein Sci.* **1998**, *7*, 1939−1946.
(132) Lombardi, A.; Marasco, D.; Maglio, O.; Di Costanzo, L.; Nasti, F.; Pavone, V. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 11922−11927.
(133) Benson, D. E.; Wisz, M. S.; Liu, W. T.; Hellinga, H. W. *Biochemistry* **1998**, *37*, 7070−7076.
(134) Nanda, V.; Rosenblatt, M. M.; Osyczka, A.; Kono, H.; Getahun, Z.; Dutton, P. L.; Saven, J. G.; DeGrado, W. F. *J. Am. Chem. Soc.* **2005**, *127*, 5804−5805.
(135) Newman, D. J.; Postgate, J. R. *Eur. J. Biochem.* **1968**, *7*, 45−50.
(136) Good, M.; Vasák, M. *Biochemistry* **1986**, *25*, 8353−8356.

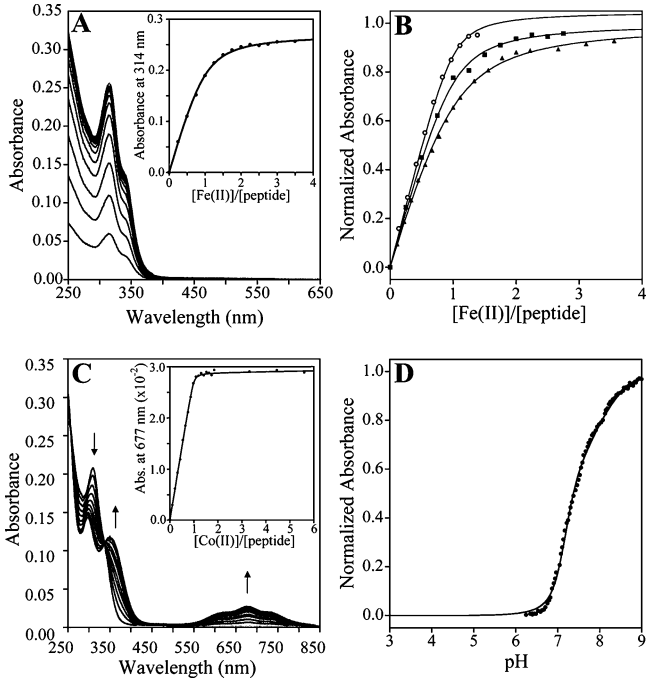


Figure 5. Determination of the **IGA** affinity for Fe(II). (A) Titration of $(NH_4)_2Fe(SO_4)_2$ into 70 μM **IGA** at pH 7.3. The increase in that absorbance at 314 nm as a function of Fe(II) added is fit to a 1:1 metal−peptide ligand model with a dissociation constant of 6 μM in the inset. (B) Influence of the solution pH on the dissociation constant of Fe^{II}-**IGA**. Titration data at pH values of 6.8 (▲), 7.3 (■), and 7.4 (○) are fit to K_d values of 40 μM, 6 μM, and 830 nM, respectively. (C) Determination of the **IGA** competition constant between Fe(II) and Co(II), $K_{comp}^{Fe/Co}$, by the addition of aqueous $CoCl_2$ into a 52.6 μM sample of Fe^{II}-**IGA** at pH 7.5. UV−visible spectra show conversion of Fe^{II}-**IGA** into Co^{II}-**IGA** LMCT with an isosbestic point at 331 nm. The inset shows the absorbance at 677 nm versus added Co(II) fit to a competition constant, $K_{comp}^{Fe/Co}$, value of 330 at pH 7.5. (D) Determination of the effective pK_a values of the cysteine ligands in Fe^{II}-**IGA**. Titration of acid into a solution of 152 μM Fe^{II}-**IGA** results in loss of the S−Fe^{II} LMCT bands. The data for Fe^{II}-**IGA** are fit to a proton competition model involving the protonation of one cysteine at pK_{a1}^{eff} of 7.4 followed by the protonation of the three remaining cysteines with a single pK_{a2}^{eff} value of 7.0. All experiments were conducted in a 20 mM HEPES buffer containing 100 mM KCl.

values. In addition, it provides the thermodynamic stability of the Fe^{II}-**IGA** complex, i.e., the formation constant of the reaction in eq 2, or K_f^{ML} in eq 3. In this system, K_f^{ML} is equivalent to β_{110}, where β_{mlh} represents the equilibrium constant for the formation of a complex containing *m* metals, *l* ligands, and *h* protons.[137]

$$\mathbf{IGA} + Fe^{II}(H_2O)_6 \rightleftharpoons Fe^{II}\text{-}\mathbf{IGA} + 6H_2O \quad (2)$$

$$K_f^{ML} = \beta_{110} = [Fe^{II}\text{-}\mathbf{IGA}]/\{[\mathbf{IGA}][Fe^{II}]\} \quad (3)$$

$$K_d = \{[Fe^{II}][\mathbf{IGA}]\}/[Fe^{II}\text{-}\mathbf{IGA}] \quad (4)$$

Absorption spectroscopy was used to measure the conditional K_d values, eq 4, for the Fe^{II}-**IGA** complex over the pH range of 6.8−7.5. As shown in Figure 5A, direct titration of $(NH_4)_2Fe(SO_4)_2$ into a 70 μM solution of **IGA** at pH 7.3 under strictly anaerobic conditions yields the Fe^{II}-**IGA** complex as demonstrated by the increase in the charge-transfer band at 314 nm. The absorption data are well fit by a 1:1 metal−peptide binding equation with a K_d value of 6 μM at pH 7.3. Figure 5B shows the determination of the Fe^{II}-**IGA** conditional dissociation constants over the pH range of 6.8−7.4. As expected, because of proton competition, the observed K_d values weaken as the pH is lowered, 830 nM (pH 7.4), 6 μM (pH 7.3), and 40 μM (pH 6.8). The Fe^{II}-**IGA** K_d value at pH 7.0, 15 μM, is quite close to that reported for the His_2Cys_2 site in CP-1, 2.5 μM at pH 7.0, despite the changes in the primary coordination sphere and global fold between these proteins.[55]

The affinity of the **IGA** ligand for Fe(II) was also determined using a competition titration with Co(II) under anaerobic conditions. The equilibrium is shown in eq 5 with the competition constant, $K_{comp}^{Fe/Co}$, defined in eq 6. Figure 5C

$$Fe^{II}\text{-}\mathbf{IGA} + \{Co^{II}(H_2O)_6\}^{2+} \rightleftharpoons Co^{II}\text{-}\mathbf{IGA} + \{Fe^{II}(H_2O)_6\}^{2+} \quad (5)$$

$$K_{comp}^{Fe/Co} = K_d^{Co}/K_d^{Fe} = \{[Fe^{II}\text{-}\mathbf{IGA}][Co^{II}]\}/\{[Fe^{II}][Co^{II}\text{-}\mathbf{IGA}]\} \quad (6)$$

shows the UV−visible spectra for the Co(II) competition titration with Fe^{II}-**IGA** at pH 7.5. The displacement of Fe(II) in **IGA** by Co(II) is clearly demonstrated by the decrease of the Fe^{II}-**IGA** absorption bands and the concomitant increase in the Co^{II}-**IGA** spectrum with an isosbestic point at 332 nm. Analysis of the spectrum, shown in the inset, yields a $K_{comp}^{Fe/Co}$ value of 330, from which a $K_d^{Fe(II)}$ value of 330 nM is determined as a result of the 1 nM value of $K_d^{Co(II)}$ at pH 7.5 (*vide infra*). On the basis of these dissociation constants, the Co^{II}-**IGA** complex exhibits 3.4 kcal/mol increased stability over the Fe^{II}-**IGA** complex at pH 7.5.

The stability of the Fe^{II}-**IGA** complex with respect to acid was also measured in order to compare it with the stability of the [4Fe−4S]-**IGA** complex (*vide infra*) because both are potential products of the self-assembly reaction used to generate the ferredoxin maquette.[93] Anaerobic titration of the Fe^{II}-**IGA** complex with dilute acid results in a bleaching of the UV−visible spectrum due to loss of the S → Fe(II) charge-transfer transitions consistent with a shift to the left of the equilibrium in eq 7. The simplest, chemically

$$\mathbf{IGA}\text{-}4H^+ + \{Fe^{II}(H_2O)_6\}^{2+} \rightleftharpoons Fe^{II}\text{-}\mathbf{IGA} + 4H^+ + 6H_2O \quad (7)$$

reasonable model that fits the data in Figure 5D involves protonation of one cysteine of Fe^{II}-**IGA** at a pK_{a1}^{eff} value of 7.4, followed by a cooperative protonation of the other three metal-bound cysteines at a pK_{a2}^{eff} value of 7.0, as shown in eqs 8 and 9. This 1:3 protonation model appears to reflect

$$pK_{a1}^{eff} \quad Fe^{II}\text{-}\mathbf{IGA}\text{-}H^+ \rightleftharpoons Fe^{II}\text{-}\mathbf{IGA} + H^+ \quad (8)$$

$$pK_{a2}^{eff} \quad Fe^{II}\text{-}\mathbf{IGA}\text{-}4H^+ \rightleftharpoons Fe^{II}\text{-}\mathbf{IGA}\text{-}H^+ + 3H^+ \quad (9)$$

the design of the ferredoxin maquette ligand, which includes three cysteines from the consensus motif plus an additional cysteine.

(137) Martell, A. E.; Motekaitis, R. J. *Determination and Use of Stability Constants*; VCH Publishers: New York, 1992.

Figure 6 shows a plot of all of the measured conditional K_d values for Fe^{II}-**IGA** as a function of the solution pH. The data are fit to an equilibrium binding model, eq 10, based on the observed 1:3 proton binding events of Fe^{II}-**IGA** and include terms for free ligand and metal−ligand complex speciation, eqs 11−14 and 15−18, respectively. (derivation in the Supporting Information).

$$-\log K_d = -\log\{[(\alpha_{\mathbf{IGA}-4H^+}(10^{-pK_a+pK_{a1}^{eff}})([10^{-3pK_a+3pK_{a2}^{eff}}]K_f^{ML})/\alpha_{M^{II}-\mathbf{IGA}-4H^+}] + [\alpha_{\mathbf{IGA}-H^+}(10^{-pK_a+pK_{a1}^{eff}})K_f^{ML})/\alpha_{M^{II}-\mathbf{IGA}-H^+}] + (\alpha_{\mathbf{IGA}}K_f^{ML}/\alpha_{M^{II}-\mathbf{IGA}})\} \quad (10)$$

$$\alpha_{\mathbf{IGA}-4H^+} = 10^{-4pH}/\Sigma_L \quad (11)$$

$$\alpha_{\mathbf{IGA}-H^+} = 10^{-pH-3pK_a}/\Sigma_L \quad (12)$$

$$\alpha_{\mathbf{IGA}} = 10^{-3pK_a-pK_a}/\Sigma_L \quad (13)$$

$$\Sigma_L = 10^{-4pH} + 10^{-pH-3pK_a} + 10^{-3pK_a-pK_a} \quad (14)$$

$$\alpha_{M^{II}-\mathbf{IGA}-4H^+} = 10^{-4pH}/\Sigma_M \quad (15)$$

$$\alpha_{M^{II}-\mathbf{IGA}-H^+} = 10^{-3pH-pK_{a2}^{eff}}/\Sigma_M \quad (16)$$

$$\alpha_{M^{II}-\mathbf{IGA}} = 10^{-3pK_{a2}^{eff}-pK_{a1}^{eff}}/\Sigma_M \quad (17)$$

$$\Sigma_M = 10^{-4pH} + 10^{-pH-3pK_{a2}^{eff}} + 10^{(-3pK_{a2}^{eff}-pK_{a1}^{eff})} \quad (18)$$

In this series of equations, eqs 10−18, the free ligand pK_a is set to the free cysteine value of 8.3, the pK_{a1}^{eff} and pK_{a2}^{eff} values are derived from direct proton competition experiments (Figure 5D), $\alpha_{\mathbf{IGA}-xH^+}$ represents the mole fraction of each free ligand protonation state, $\alpha_{M^{II}-\mathbf{IGA}-xH^+}$ is the mole fraction of each metal−ligand protonation state, and Σ_L and Σ_M are the total ligand and metal−ligand concentrations, respectively. It is worth noting that the differences between the acid dissociation constants of the free ligand, pK_a, and metal−ligand complex, pK_a^{eff}, dictate the value of the conditional formation constant K_d at any solution pH value.

The Fe^{II}-**IGA** data are well described by eq 10, and the formation constant value, K_f^{ML}, derived from this analysis is 5.0×10^8 M^{-1}, which indicates that the $Fe^{II}(S-Cys)_4$ site contributes −11.8 kcal/mol to metallopeptide stability. Another way to express the effectiveness of a chelator for a metal ion is to calculate the pM value,[138,139] or the negative log of the concentration of free metal present in a solution of 1 μM metal and 10 μm ligand at pH 7.4. As given in Table 2, the pM value for Fe^{II}-**IGA** is 7.0.

While Fe^{II}-**IGA** is a good spectroscopic model for a reduced rubredoxin active site, its weak Fe(II) binding thermodynamics highlight a significant issue in the design of synthetic rubredoxins. Namely, the Fe^{II}-**IGA** complex is in equilibrium with complexes where at least one cysteinate ligand is protonated at physiological pH values, as shown in eq 8. For example, the data indicate that a 1:1 mixture of ferrous ion and the **IGA** ligand at pH 7.0 will generate only 17% of the desired Fe^{II}-**IGA** complex, with the remainder of the ligand equally divided between Fe^{II}-**IGA**-H^+ and Fe^{II}-**IGA**-$4H^+$ species.

The weak dissociation constant for the Fe^{II}-**IGA** complex and its proton sensitivity near physiological pH values is

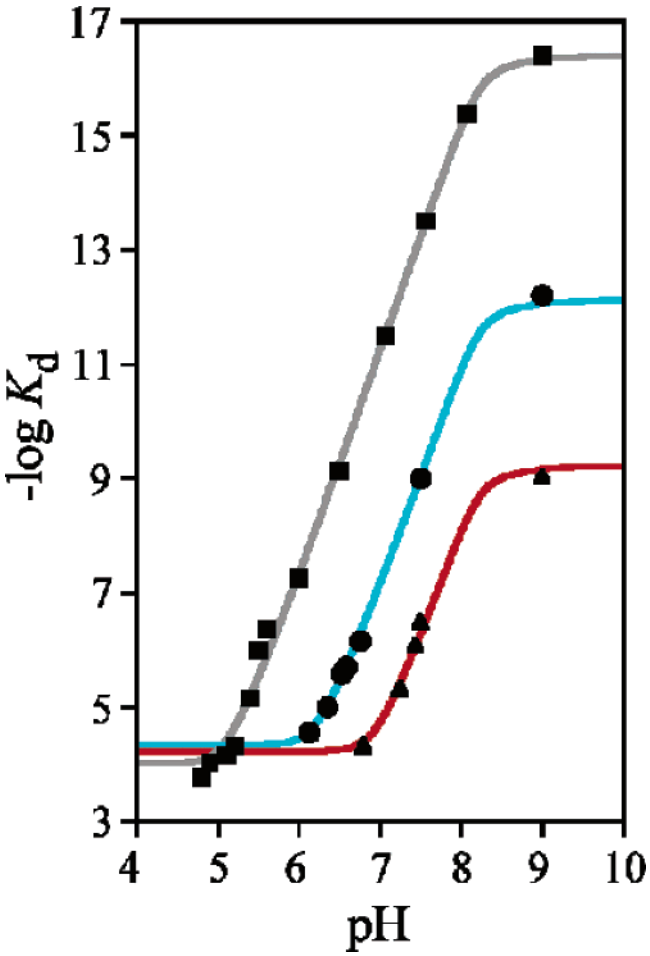


Figure 6. pH dependence of the conditional dissociation constant values for Fe^{II}-**IGA** (▲), Co^{II}-**IGA** (■), and Zn^{II}-**IGA** (■). Individual K_d values are plotted versus the solution pH. The pH dependence of the K_d data for each metal is fit using eq 10 in the text. Each fit is based on the pK_a value of the ligand cysteines and set to the free cysteine value of 8.3, and the effective proton dissociation constant values of the metal−ligand complexes, pK_{a1}^{eff} and pK_{a2}^{eff}, were determined by proton competition titrations, as shown in Figures 5D, 7D, and 8D. Each K_d value was determined using UV−visible or fluorescence spectroscopy by either direct or competition titration (metal ion or EDTA) at various pH values as detailed in the text and Figures 5B, 7B, and 8B.

Table 2. Thermodynamic Stabilities of the Fe(II), Co(II), and Zn(II) Complexes of **IGA** (t = 25° C, I = 0.1 M KCl)

complex	K_d (pH 7.0)	pK_{a1}^{eff}	pK_{a2}^{eff}	K_f^{ML} [a] (M^{-1})	log β_{114} [b] (M^{-5})	pM [c]
Fe^{II}-**IGA**	15 μM	7.4	7.0	5.0×10^8	37.2	7.0
Co^{II}-**IGA**	80 nM	6.7	6.1	4.2×10^{11}	36.6	9.6
Zn^{II}-**IGA**	3 pM	5.5	5.1	8.0×10^{15}	36.7	13.7

[a] $K_f^{ML} = \beta_{110}$. [b] $\beta_{114} = K_f^{ML}(1/K_{a1}^{eff})(1/K_{a2}^{eff})^3$. [c] pM = −log [$M_{free}$] at 1 μM [M_{total}], 10 μM [L_{total}], at pH 7.4.[138]

(138) Harris, W. R.; Carrano, C. J.; Raymond, K. N. *J. Am. Chem. Soc.* **1979**, *101*, 2213−2214.
(139) Garrett, T. M.; Miller, P. W.; Raymond, K. N. *Inorg. Chem.* **1989**, *28*, 128−133.

likely common in most designed $Fe^{II}(S-Cys)_4$ proteins where protein folding does not contribute to metalloprotein stability.[80,131−134] Furthermore, this issue clearly illustrates the necessity for evaluating the pH stability of designed metalloproteins, especially in the case of thiolate−peptide complexes of ferrous ion. The observed weak affinity of Fe(II) for thiolate ligands in competition with water may be a critical factor in the fact that, to date, designed rubredoxin proteins have met with limited success in replicating the rubredoxin Fe(II)/Fe(III) redox cycling function even in the presence of exogenous thiols.[133,134]

Thermodynamic Analysis of the Co(II) Affinity of IGA

To fully describe the thermodynamics of Co^{II}-**IGA** formation, direct metal-ion titrations, competition experiments with Fe(II), Zn(II), and H^+ were performed using absorption and fluorescence spectroscopies. As with the Fe^{II}-**IGA** complex, coupling the various data sets provides a full description of the coordination equilibria involved in complex formation and a determination of the formation constant, K_f^{ML}, for Co^{II}-**IGA**. While the stability of the Fe^{II}-**IGA** complex is limited, the Co^{II}-**IGA** complex shows considerable thermodynamic stability in an aqueous solution.

Figure 4B shows the UV−visible spectrum of Co^{II}-**IGA**, which is consistent with a tetrahedral tetrathiolate coordination geometry at the metal center. The intensities of the ligand-field bands at 630 nm ($\epsilon = 400\ M^{-1}\ cm^{-1}$), 686 nm ($\epsilon = 570\ M^{-1}\ cm^{-1}$), and 728 nm ($\epsilon = 540\ M^{-1}\ cm^{-1}$) demonstrate a pseudotetrahedral primary coordination sphere. The intensity ($\epsilon = 3800\ M^{-1}\ cm^{-1}$) of the LMCT band at 304 nm is indicative of four Co^{II}−S bonds.[140] Detailed EXAFS analysis demonstrated a $Co^{II}(S-Cys)_4$ primary coordination sphere in Co^{II}-**IGA** with an average Co^{II}−S bond length of 2.31 Å.[80] These spectroscopic and metrical parameters are similar to those of structurally characterized small-molecule models[141,142] such as $\{Co^{II}(ethanedithiolate)_2\}^{2-}$ and $\{Co^{II}(S-Ph)_4\}^{2-}$ as well as Co(II)-substituted natural proteins with $M^{II}(S-Cys)_4$ active sites such as rubredoxin[143] and the GATA zinc finger protein.[144]

The conditional K_d values for Co^{II}-**IGA** were determined over the pH range of 6.1−7.0 using absorption spectroscopy. As shown in Figure 7A, titration of $CoCl_2$ into a 50 μM solution of **IGA** at pH 6.5 leads to formation of the Co^{II}-**IGA** complex, as evidenced by the increase in the charge-transfer and ligand-field bands. Figure 7B shows the conditional dissociation constants for Co^{II}-**IGA** determined over the pH range of 6.1−7.0. The K_d values determined at each pH are 28 μM (pH 6.1), 10 μM (pH 6.2), 2 μM (pH 6.5), and tighter than 0.5 μM (pH 7.0). In addition, a K_d value of 1.0 nM was accurately determined at pH 7.5 using the increase in the tryptophan fluorescence intensity due to the loss of quenching observed upon substitution of the bound high-spin Co(II) ($S = {}^3/_2$) for Zn(II) in **IGA**.

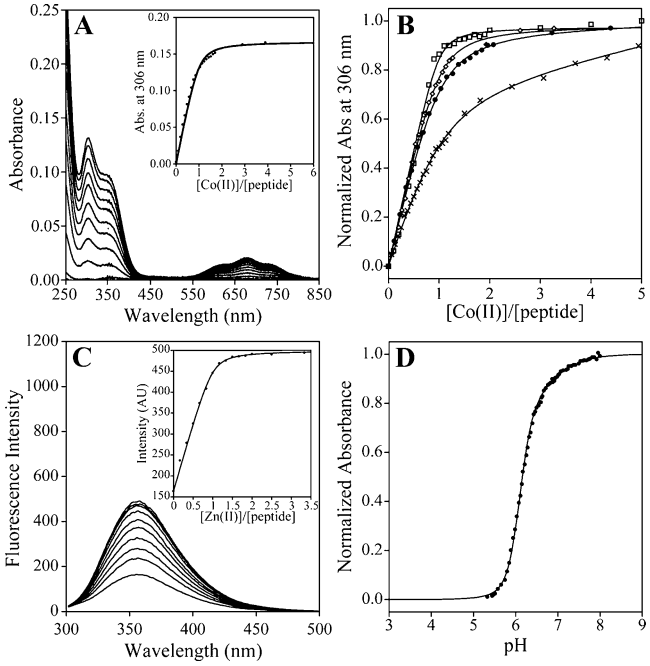


Figure 7. Determination of the **IGA** affinity for Co(II). (A) Titration of $CoCl_2$ into 55 μM **IGA** at pH 6.5 in 20 mM MES buffer containing 100 mM KCl. The increase in absorbance at 306 nm as a function of Co(II) added is fit to a 1:1 metal−peptide ligand model with a dissociation constant of 13.5 μM in the inset. (B) Effect of the solution pH on the dissociation constant of Co^{II}-**IGA**. Direct Co(II) titrations into **IGA** at pH values of 6.1 (×), 6.2 (●), 6.5 (○), and 7.0 (□) are shown fit to dissociation constant values of 28 μM, 10 μM, 2 μM, and tighter than 500 nm, respectively. (C) Determination of the **IGA** competition constant between Co(II) and Zn(II), $K_{comp}^{Co/Zn}$, by the addition of aqueous $ZnCl_2$ into a 1.0 μM sample of Co^{II}-**IGA** at pH 6.5 containing 200 equiv of Co(II). The fit demonstrates a competition constant, $K_{comp}^{Co/Zn}$, value of 6700 at pH 6.5 as shown in the inset. (D) Determination of the effective pK_a values of the cysteine ligands in Co^{II}-**IGA**. Titration of acid into a solution of 132 μM Co^{II}-**IGA** results in a bleaching of the optical spectrum. The data for Co^{II}-**IGA** are shown fit to a proton competition model involving the protonation of one cysteine at pK_{a1}^{eff} of 6.7 followed by the protonation of the three remaining cysteines with a single pK_{a2}^{eff} value of 6.1.

A competition constant for the displacement of Co(II) in Co^{II}-**IGA** by Zn(II) was determined using fluorescence spectroscopy in the presence of a 200-fold excess of $\{Co^{II}(H_2O)_6\}^{2+}$, as shown in eq 19. Figure 7C shows the

$$Co^{II}\text{-}\mathbf{IGA} + \{Zn^{II}(H_2O)_6\}^{2+} \rightleftharpoons Zn^{II}\text{-}\mathbf{IGA} + \{Co^{II}(H_2O)_6\}^{2+} \quad (19)$$

fluorescence emission spectra of the Zn(II) competition titration with Co^{II}-**IGA** at pH 6.5. The formation of Zn^{II}-**IGA** results in an increase in the fluorescence emission intensity due to the substitution of the paramagnetic Co(II) by diamagnetic Zn(II). A $K_{comp}^{Co/Zn}$ value of 6700 was determined from a fit to the data, as shown in the inset of Figure 7C. Thus, the Zn^{II}-**IGA** complex exhibits 5.2 kcal/mol increased stability over the Co^{II}-**IGA** complex at pH 6.5.

The Co^{II}-**IGA** complex was also titrated with dilute HCl under anaerobic conditions, as in the Fe^{II}-**IGA** case. As the pH is lowered, the S → Co(II) charge-transfer and ligand-field bands lose intensity, consistent with protonation and

(140) Lever, A. B. P. *Inorganic Electronic Spectroscopy*; Elsevier: Amsterdam, The Netherlands, 1984.
(141) Dance, I. G. *J. Am. Chem. Soc.* **1979**, *101*, 6264−6273.
(142) Rao, C. P.; Dorfman, J. R.; Holm, R. H. *Inorg. Chem.* **1986**, *25*, 428−439.
(143) May, S. W.; Kuo, J. Y. *Biochemistry* **1978**, *17*, 3333−3338.
(144) Ghering, A. B.; Shokes, J. E.; Scott, R. A.; Omichinski, J. G.; Godwin, H. A. *Biochemistry* **2004**, *43*, 8346−8355.

decomplexation of tetrahedral Co^{II}-**IGA** to yield octahedral $\{Co^{II}(H_2O)_6\}^{2+}$ with Laporte-forbidden ligand-field bands. The data shown in Figure 7D are best fit to two effective pK_a values weighted by the number of protons, one and three, as observed for Fe^{II}-**IGA**. The measured pK_{a1}^{eff} value is 6.7 ($1H^+$), and the value for pK_{a2}^{eff} is 6.1 ($3H^+$). The more acidic pK_a^{eff} values of Co^{II}-**IGA** compared to Fe^{II}-**IGA** are fully consistent with the increased affinity of the ligand for Co(II) over Fe(II) and indicate a 4.0 kcal/mol difference in their respective formation constants.

Figure 6 shows the pH dependence of the conditional K_d values for Co^{II}-**IGA** fit to eq 10. As with the corresponding analysis of Fe^{II}-**IGA**, the data are fit to an equilibrium binding model based on the 1:3 proton binding observed for Co^{II}-**IGA** with pK_{a1}^{eff} and pK_{a2}^{eff} values of 6.7 and 6.1, respectively. These pK_a^{eff} values indicate that there is substantial protonation of the metal−ligand complex under physiological conditions; i.e., 38.6% exists as the Co^{II}-**IGA**-H^+ species at pH 7.0. At pH values where the **IGA** ligand is fully deprotonated, a K_f^{ML} value of 4.2×10^{11} M^{-1} is extrapolated based on the fit shown in Figure 6. The pH modulation of this K_f^{ML} value yields a pM value of 9.6, as listed in Table 2.

Thermodynamic Analysis of the Zn(II) Affinity of IGA

The **IGA** ligand is an avaricious binder of Zn(II) in an aqueous solution. Direct metal-ion titrations and proton competition experiments were performed using fluorescence spectroscopy to determine the thermodynamic stability of Zn^{II}-**IGA** under a variety of solution conditions. Analysis of the data sets provides a detailed description of the coordination equilibria and the value of the Zn^{II}-**IGA** formation constant, K_f^{ML}. These data demonstrate that **IGA** has a femtomolar affinity for Zn(II), which rivals natural Zn(II) proteins and enzymes.

Fluorescence spectroscopy was used to determine the conditional dissociation constants of the Zn^{II}-**IGA** complex over the pH range of 4.8−8.0. Figure 8A shows that the binding of Zn(II) to **IGA** under anaerobic conditions at pH 5.5 results in an increase in the fluorescence emission intensity of the C-terminal tryptophan. The observed rise in tryptophan emission at 357 nm ($\lambda^{ex} = 280$ nm) is due to decreased fluorescence quenching by cysteine thiols in the free ligand as Zn(II) binds.[145] Analysis of the data using a 1:1 metal−ligand equilibrium binding model manifests a conditional dissociation constant value of 1 μM at pH 5.5. Figure 8B shows the determination of the K_d value of Zn^{II}-**IGA** over the pH range of 5.0−6.5, where direct metal-ion titrations were used. The K_d values measured were 100 μM (pH 5.0), 50 μM (pH 5.3), 1 μM (pH 5.5), and tighter than 1 nM (pH 6.5). Over the pH range of 7.0−8.0, the determination of the Zn^{II}-**IGA** conditional K_d values necessitated the use of a competing chelator.[146] Ethylenediaminetetraacetic acid (EDTA) was chosen for its tight Zn(II)

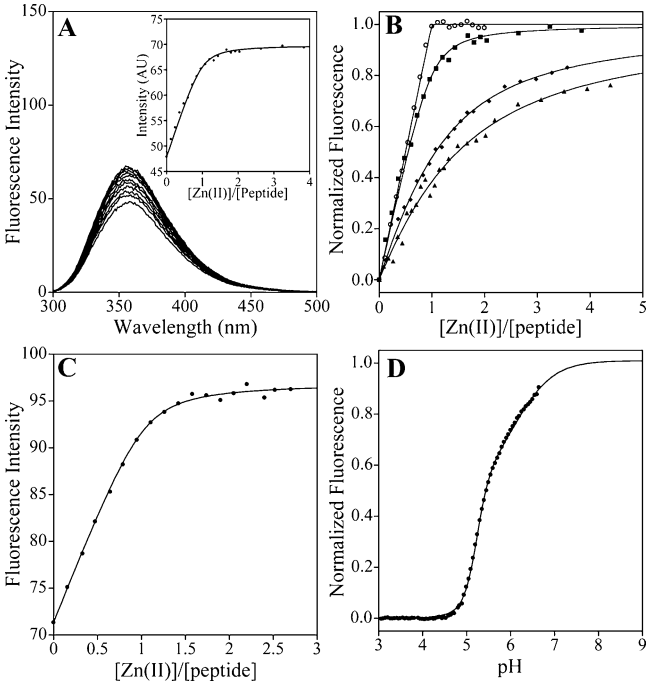


Figure 8. Determination of the **IGA** affinity for Zn(II). (A) Titration of $ZnCl_2$ into 22 μM **IGA** at pH 5.5 in a 20 mM MES buffer with 100 mM KCl. The increased fluorescence emission at 357 nm as a function of Zn(II) added is fit to a 1:1 metal−peptide ligand model with a dissociation constant of 1.0 μM in the inset. (B) Influence of the solution pH on the dissociation constant of Zn^{II}-**IGA** measured by fluorescence. Titration data at pH values of 5.0 (▲), 5.3 (◆), 5.5 (■), and 6.0 (○) are fit to K_d values of 100 μM, 50 μM, 1 μM, and 50 nM, respectively. (C) Determination of the K_d value of Zn^{II}-**IGA** at pH 8.0 in 20 mM KPi and 100 mM KCl. The fluorescence emission intensity was monitored as $ZnCl_2$ was titrated into a 13 μM solution of **IGA** containing 13 μM EDTA and fit to a competition constant of 18, which yields a 0.4 fM K_d value of Zn-**IGA** at pH 8.0. (D) Determination of the effective pK_a values of the cysteine ligands in Zn^{II}-**IGA**. Titration of acid into a solution of 60 μM Zn^{II}-**IGA** results in a decrease in tryptophan fluorescence due to quenching by free thiols. The data for Zn^{II}-**IGA** are best fit to a proton competition model involving the protonation of one cysteine at pK_{a1}^{eff} of 5.5 followed by the protonation of the three remaining cysteines with a single pK_{a2}^{eff} value of 5.1.

affinity, K_f^{ML} value of 3×10^{16} M^{-1}.[147] As shown in Figure 8C, titration of Zn(II) into a solution containing 13 μM **IGA** and 13 μM EDTA at pH 8.0 yields a conditional K_d value of 0.4 fM. At pH 7.0, the K_d value of Zn^{II}-**IGA** was determined to be 4 pM, a value nearly identical with that of the designed Cys_4 site in CP1−CCCC with a K_d of 1.1 pM[55b] and well within the range observed for natural zinc proteins. The Zn(II) affinity of **IGA** at pH 7.0 is about equal to those of human carbonic anhydrase II (4 pM at pH 7.0)[148] and metallothionein (0.1 pM at pH 7.0)[149] and weaker than those of the zinc sensor proteins ZntR (1.5 fM at pH 7.0)[127a] and Zur (1.1 fM at pH 7.6).[127b]

The direct pH titration of Zn^{II}-**IGA** was measured using fluorescence spectroscopy. The data in Figure 8D are fit to pK_{a1}^{eff} and pK_{a2}^{eff} values of 5.5 ($1H^+$) and 5.1 ($3H^+$), respectively. This is consistent with the enhanced basicity of one cysteine over the remaining three, as seen in the Fe^{II}- and Co^{II}-**IGA** complexes. The enhanced thermodynamic basicity

(145) Harris, D. L.; Hudson, B. S. *Biochemistry* **1990**, *29*, 5276−5285.
(146) EDTA competition experiments were also used to measure the Co(II) and Fe(II) K_d values at pH 9.0.
(147) Martell, A. E.; Smith R. M. *Critical Stability Constants*; Plenum Press: New York, 1974; Vol. 1.
(148) McCall, K. A.; Fierke, C. A. *Biochemistry* **2004**, *43*, 3979−3986.
(149) Maret, W. *J. Nutr.* **2003**, *133*, 1460s−1462s.

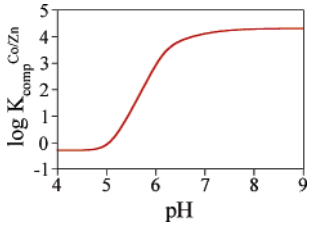


Figure 9. Simulation of the pH dependence of the competition constant for **IGA** binding Zn(II) over Co(II), $K_{comp}^{Co/Zn}$. The curve was derived from the ratio of the Zn(II) and Co(II) fits in Figure 6.

of one cysteinate may translate into enhanced kinetic nucleophilicity and provide functional models of the Zn^{II}-$(S-Cys)_4$ site in O^6-methylguanine transferases.[41] Figure 6 shows a fit of the pH dependence of the K_d values to eq 10, which indicates a Zn^{II}-**IGA** formation constant, K_f^{ML}, value of 8.0×10^{15} M^{-1} and a pM value of 13.7. Therefore, the Zn(II) complex of **IGA** is 5.8 and 9.8 kcal/mol more stable than the corresponding Co(II) and Fe(II) complexes, respectively.

The Zn^{II}-**IGA** K_d value of 0.4 fM at pH 8.0 represents the tightest Zn(II) affinity measured to date for a designed protein system and supersedes the values reported for natural Zn(II) proteins including the 1.4-fM value of the Zn(II) sensor/regulator protein, ZntR, at pH 8.0.[127] The affinity of **IGA** for Zn(II) is significantly modulated at pH values below 8.5 by proton competition. As demonstrated by the gray curve in Figure 6, the conditional dissociation constant for Zn(II) has a $[H^+]^4$ dependence near physiological pH. Notably, at pH 7.0, the K_d value of Zn^{II}-**IGA** is weaker than that of ZntR because of the differences in their ligand pK_a values. ZntR shows no pH dependence, while Zn^{II}-**IGA** demonstrates a $[H^+]^4$ modulation over the pH range of 7.0−8.0, demonstrating that the ligand pK_a values of ZntR are lower than 7.0. These data highlight the difficulty in comparing conditional K_d values without a full description of the underlying protonation equilibria.[54]

In the case of Zn^{II}- and Co^{II}-**IGA**, the conditional competition constant between these two metals is highly pH-dependent because of differences in the pK_a^{eff} values since the free ligand pK_a values are identical. Figure 9 shows how the value of the conditional $K_{comp}^{Co/Zn}$ varies with the solution pH based on the ratio of the equations used to fit the Co^{II}- and Zn^{II}-**IGA** data sets in Figure 6. This variation reflects differences in the underlying proton dissociation constants of the metal−ligand species, as given in Table 2. The inherent $K_{comp}^{Co/Zn}$ value of 2.3×10^4 or 5.8 kcal/mol is only observed at high pH values, where all of the ligand and metal−ligand complexes are fully deprotonated, i.e., above pH 9. As the pH is lowered, the protonation of the Co^{II}-**IGA** complex begins to attenuate the observed values of $K_{comp}^{Co/Zn}$. The conditional $K_{comp}^{Co/Zn}$ value drops to a value of 870, 4.0 kcal/mol, at pH 6.0. At this pH, Co^{II}-**IGA** exists in an equilibrium between the unprotonated (3%), monoprotonated (19.5%), and tetraprotonated (77.5%) forms, while Zn^{II}-**IGA** is present as the unprotonated (76%) and monoprotonated (24%) forms. While comparison of conditional competition constants between Zn(II) and Co(II) is common in the zinc protein literature,[19,132,150,151] this analysis highlights the inaccuracy obtained without a full description of the principal proton dissociation constants of the metal−ligand species.[152]

There has been considerable debate in the literature concerning the protonation state of cysteine ligands bound to Zn(II) because the resolution of most Zn(II) protein structures in the PDB is not sufficient to discern the metal−ligand bond elongation expected upon thiolate protonation.[23,153−156] The coordination equilibria analysis presented herein provides a methodology for determining the protonation state of ligands bound to metals in metalloproteins. In the case of Zn^{II}-**IGA** at pH 7.0, we observe only 3% protonation of one cysteine thiolate in this solvent-exposed $Zn^{II}(S-Cys)_4$ site. The data for Fe^{II}-**IGA** evince that mixtures of metal−ligand proton species exist at pH 7.0 because of the 9.8 kcal/mol weaker inherent affinity for Fe(II) relative to Zn(II). Thus, this analysis highlights that ligand protonation is more likely to occur in $Fe^{II}(S-Cys)_4$ proteins at physiological pH.

Berg has argued that zinc finger proteins prefer Zn(II) over Co(II) because of the differential loss of LFSE upon conversion of an aqueous octahedral M(II) ion into a protein-bound tetrahedral M(II) ion.[54] The experimentally observed 5.8 kcal/mol enhanced inherent affinity of **IGA** for Zn(II) over Co(II) is close to the 6.6 kcal/mol value for the change in LFSE estimated based on a $Co^{II}(S-Cys)_4$ tetrahedral ligand-field splitting value, Δ_t, of 4295 cm^{-1}.[49,157] However, this LFSE argument fails to account for the 4.0 kcal/mol *lower* stability of Fe^{II}-**IGA** relative to Co^{II}-**IGA** because it predicts the Fe(II) peptide complex to be at least 2.7 kcal/mol *more* stable than the Co(II) peptide. Thus, LFSE is not the predominant contributor to the observed 9.8 kcal/mol stabilization of Zn(II) over Fe(II) in **IGA**, and the correlation of the spectroscopic value of the Co(II) LFSE and the experimental $K_{comp}^{Co/Zn}$ value may be coincidental.

(150) (a) Bombarda, E.; Cherradi, H.; Morellet, N.; Roques, B. P.; Mély, Y. *Biochemistry* **2002**, *41*, 4312−4320. (b) Bombarda, E.; Morellet, N.; Cherradi, H.; Spiess, B.; Bouaziz, S.; Grell, E.; Roques, B. P.; Mély, Y. *J. Mol. Biol.* **2001**, *310*, 659−672. (c) Mély, Y.; De Rocquigny, H.; Morellet, N.; Roques, B. P.; Gérard, D. *Biochemistry* **1996**, *35*, 5175−5182.

(151) (a) Bavoso, A.; Ostuni, A.; Battistuzzi, G.; Menabue, L.; Saladini, M.; Sola, M. *Biochem. Biophys. Res. Commun.* **1998**, *242*, 385−389. (b) Posewitz, M. C.; Wilcox, D. E. *Chem. Res. Toxicol.* **1995**, *8*, 1020−1028. (c) Predki, P. F.; Sarkar, B. *J. Biol. Chem.* **1992**, *267*, 5842−5846. (d) Borden, K. L. B.; Martin, S. R.; O'Reilly, N. J.; Lally, J. M.; Reddy, B. A.; Etkin, L. D.; Freemont, P. S. *FEBS* **1993**, *335*, 255−260. (e) Benson, D. E.; Wisz, M. S.; Liu, W.; Hellinga, H. W. *Biochemistry* **1998**, *37*, 7070−7076.

(152) Lachenmann, M. J.; Ladbury, J. E.; Dong, J.; Huang, K.; Carey, P.; Weiss, M. A. *Biochemistry* **2004**, *43*, 13910−13925.

(153) Clark-Baldwin, K.; Tierney, D. L.; Govindaswamy, N.; Gruff, E. S.; Kim, C.; Berg, J. M.; Koch, S. A.; Penner-Hahn, J. E. *J. Am. Chem. Soc.* **1998**, *120*, 8401−8409.

(154) Simonson, T.; Calimet, N. *Proteins* **2002**, *49*, 37−48.

(155) Dudev, T.; Lim, C. *J. Am. Chem. Soc.* **2002**, *124*, 6759−6766.

(156) Alberts, I. L.; Nadassy, K.; Wodak, S. J. *Protein Sci.* **1998**, *7*, 1700−1716.

(157) As noted in ref 54, moving a high-spin d^7 Co(II) from octahedral to tetrahedral symmetry results in a LFSE loss of $-{}^4/_5\Delta_0 - {}^6/_5\Delta_t$. Using a Δ_0 value of 9300 cm^{-1} and a Δ_t value of 4295 cm^{-1} (from ref 49), a LFSE loss of 6.6 kcal/mol is estimated for Co(II). For high-spin Fe(II), a LFSE loss of $-{}^2/_5\Delta_0 - {}^3/_5\Delta_t$ is expected. Using a Δ_0 value of 9400 cm^{-1} and a Δ_t values of 4000 and 5000 cm^{-1} (the lower and upper limits from ref 49), a LFSE loss between 2.2 and 3.9 kcal/mol is estimated for Fe(II).

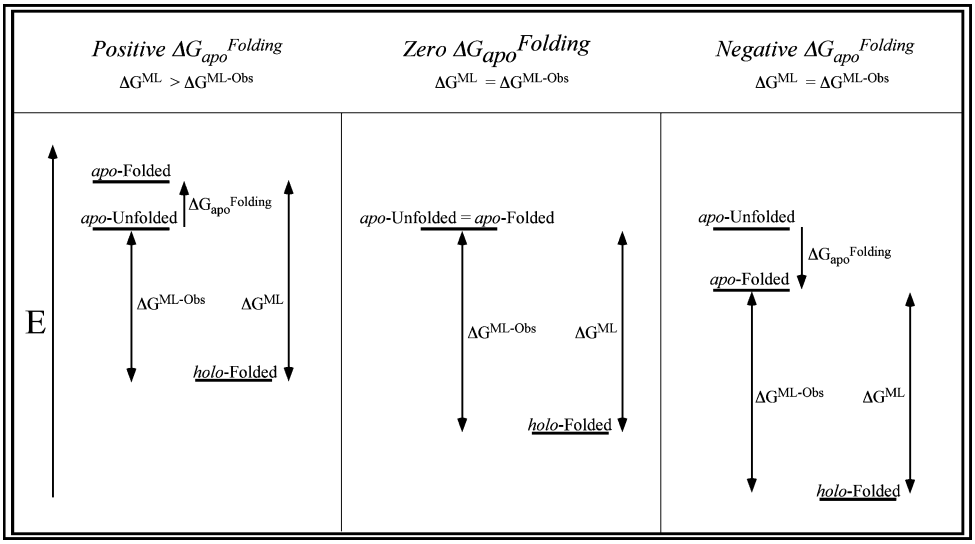


Figure 10. Comparison of the observed and maximal metal-ion affinities in various energetic scenarios for metalloprotein folding. (Left) An apo state binds the metal ion and folds into the holo-folded state with an observed metal-ion binding energy of ΔG^{ML-Obs}. This value is less than the maximal metal-ion binding energy, ΔG^{ML}, which represents the energy difference between apo-unfolded and -folded states. (Middle) In this scenario, the observed and maximal metal-ion binding energy values are equivalent because the apo-unfolded and -folded states are degenerate. (Right) The binding of a metal ion to the folded protein scaffold with a preorganized metal-ion binding site also results in a situation where ΔG^{ML-Obs} is equal to ΔG^{ML} because the apo-folded state is lower in energy than the apo-unfolded state. For clarity, the energies of the apo-unfolded states and the value of ΔG^{ML} for each panel were kept constant, and the holo-unfolded state was omitted.

The inherent Fe^{II}-, Co^{II}-, and Zn^{II}-**IGA** complex affinities are qualitatively consistent with the Irving−Williams series, which correlates the second ionization potential of the metal with the relative order of M^{II}-L complex stabilities, i.e., $Mn^{II} < Fe^{II} < Co^{II} < Ni^{II} < Zn^{II} < Cu^{II}$. Solomon and co-workers have recently shown that this trend reflects the greater covalency of M^{II}−S bonding.[158] The lack of deviation from expectations based on the Irving−Williams series or hard−soft acid−base theory is not unexpected for an unfolded peptide ligand, which does not impose stringent geometric requirements on the bound metal. Indeed, such deviations are relatively rare in designed proteins, with one notable example being the diiron DueFerro2 protein of DeGrado and co-workers.[159,160]

The inherent affinity of the Zn^{II}-**IGA** complex provides significant insight into the thermodynamics of metal-induced protein-folding studies of zinc-binding proteins. Comparison of the Zn^{II}-**IGA** data to the limited set of conditional K_d values available for natural and designed Cys_4 zinc fingers in the literature indicates that they possess very similar conditional K_d values and may possess similar values of K_f^{ML}.[19,144] In addition, the pH dependencies of Zn(II) binding to **IGA** and the His_1Cys_3 site in the C-terminal zinc finger of the HIV-1 nucleocapsid protein are virtually superimposable, demonstrating that both have similar formation constant values, K_f^{ML}.[150] Analysis of the pH dependence of Zn(II) binding to this natural zinc finger protein appears to demonstrate only a 10-fold weaker formation constant relative to Zn^{II}-**IGA**, which likely reflects the change in the primary coordination sphere, i.e., His_1Cys_3 to Cys_4. Because **IGA** is minimal in size and lacks a regular secondary structure in both the Zn(II)-bound and unbound states, we suggest that the energy required to fold this peptide is negligible and, thus, the observed −21.6 kcal/mol binding energy of **IGA** to Zn(II) represents nearly the maximal value possible for Zn(II) binding to a tetrahedral $(S-Cys)_4$ coordination sphere.

This conclusion that **IGA** binds metal ions with the maximal affinity for the ligand set assumes that the free energy difference between the apo-unfolded and -folded states, $\Delta G_{apo}^{folding}$, is zero, as shown in the center scenario of Figure 10. In this specific case, the observed metal−ligand binding energy, ΔG^{ML-Obs}, is equivalent to the maximal value, ΔG^{ML}, because the energy required to fold the apo state is zero, $\Delta G_{apo}^{folding} = 0$. The scenario on the right in Figure 10 shows that the observed and maximal metal−ligand binding energies are also equivalent in the cases where the metal ion binds to a folded *apo*-protein scaffold with its ligands optimally positioned for metal-ion binding, e.g., carbonic anhydrase. In these cases, the *apo*-protein folding energy, $\Delta G_{apo}^{folding} < 0$, does not alter ΔG^{ML-Obs}, which is equal to the maximal metal−ligand binding energy, ΔG^{ML}. The left scenario in Figure 10 shows that a value of ΔG^{ML-Obs} lower than the maximal, ΔG^{ML}, indicates that part of the metal−ligand binding energy is being used to overcome a thermodynamic barrier to protein folding, i.e., $\Delta G_{apo}^{folding} > 0$. Our conclusion that **IGA** binds metal ions with the maximal affinity for a tetrahedral tetrathiolate site indicates that **IGA** does not utilize its metal−ligand binding energy to fold the protein, consistent with its unstructured nature.

A minimal energetic difference between the apo-unfolded and -folded states suggested for **IGA** is supported by a

(158) Gorelsky, S. I.; Basumallick, L.; Vura-Weis, J.; Sarangi, R.; Hodgson, K. O.; Hedman, B.; Fujisawa, K.; Solomon, E. I. *Inorg. Chem.* **2005**, *44*, 4947−4960.
(159) Lombardi, A.; Summa, C. M.; Geremia, S.; Randaccio, L.; Pavone, V.; DeGrado, W. F. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 6298−6305.
(160) Pasternak, A.; Kaplan, S.; Lear, J. D.; DeGrado, W. F. *Protein Sci.* **2001**, *10*, 958−969.

number of lines of evidence in the zinc finger literature. First, Dahiyat and Mayo have redesigned a zinc finger protein to fold into a stable $\beta\beta\alpha$ structure without the metal or its ligands.[65a] The ability of a collection of protein−protein interactions, each relatively weak in comparison to a metal−ligand interaction, to provide enough of a thermodynamic contribution to make $\Delta G_{apo}^{folding} < 0$ is consistent with the apo-unfolded state of the original zinc finger scaffold not being high in energy. Second, Imperiali has designed a series of zinc finger protein sequences that stabilize the apo-folded state of the protein to the point where it is folded in the absence of a bound metal.[125] The sequence with the most stable apo-folded state, BBA1, shows a 7-fold, 1.1 kcal/mol, decrease in the Zn(II) affinity compared with the unfolded prototype sequence. The rigid fold of the BBA1 protein does not optimally orient the ligands for metal-ion binding and, therefore, the 7-fold decrease in Zn(II) affinity observed likely reflects the energy required to reorient the ligands for Zn(II) binding. The conformational change required to bind the metal introduces a positive ΔG term, which weakens the observed metal-ion binding energy, ΔG^{ML-Obs}, relative to ΔG^{ML}. Third, recent studies on Cu(I) binding to a peptide containing the methionine-rich MxxMxxM sequence, a Mets motif involved in metal-ion trafficking, yields a 2.5 μM conditional K_d that is as tight as the same sequence in the larger natural protein scaffold, the copper transporter Ctr1.[161] Last, a Folding@home molecular dynamics simulation result indicates that the unfolded ensemble of a $\beta\beta\alpha$ protein fold corresponds to the native folded state in the average sense, i.e., the mean structure hypothesis.[162] This finding is consistent with a limited energetic difference between the apo-unfolded and -folded states of natural zinc finger proteins and not the large differences estimated based on unfolded proteins in extended conformations.[59] Thus, this compilation of evidence suggests that it is not unlikely that the difference between the apo-unfolded and -folded states of **IGA** is zero and that it possesses the maximal metal−ligand affinity for a tetrahedral tetracysteinate binding site.

These observations that protein ligands with metal-induced protein-folding events appear to have near-maximal affinities for their ligand sets have significant implications for metalloprotein engineering. The scenarios in Figure 10 indicate that folded *apo*-protein scaffolds containing optimally preorganized metal-ion binding sites and unstructured peptide ligands can both possess the maximal metal-ion affinities for their ligand sets. Depression of the observed metal-ion affinities requires a highly destabilized apo-folded state, as shown in the left scenario of Figure 10, or a conformational change from a well-structured apo-folded state upon metal-ion binding. The latter situation is exemplified by the BBA1 design,[125] while the former is likely the case in situations where the metal ion is involved in the assembly of its binding site from multiple unfolded subunits.[132] While both protein assembly reactions and conformational changes can weaken metal-ion affinities, the magnitude of the maximal metal-ion affinity easily provides enough energy to compensate for these events and can even be used to drive a sequence from a stably folded trimeric coiled coil in the apo state to a stably folded monomeric $\beta\beta\alpha$ structure in the holo state.[163]

The maximal metal-ion binding affinity is only observed when the ligating residues of a given protein ligand are fully deprotonated. Below the pK_a values of the free ligand, the maximal affinity is attenuated by proton competition, thus imposing a pH dependence on the conditional dissociation constant, K_d. Lowering the free ligand pK_a provides a means with which to achieve proton-independent binding over a broader pH range. However, simply lowering the free ligand pK_a without lowering the metal-bound pK_a values, pK_a^{eff}, will lower the observed maximal binding affinity. Thus, modulating the ligand pK_a and metal−ligand pK_a^{eff} values provides a mechanism with which to increase the apparent stability constants under physiological conditions. For example, **IGA** has a 3-fold tighter Zn(II) affinity compared to the zinc sensor protein ZntR at pH 8.0, yet ZntR has a 5000-fold tighter Zn(II) affinity at pH 7.0.[127] The observed pH independence of the conditional dissociation constants of ZntR between pH 7.0 and 8.0 indicates that its pK_a and pK_a^{eff} values are lower than those in **IGA**. The depression of the cysteine ligand pK_a values in ZntR is likely a consequence of the preorganization of the metal-ion binding site due to protein folding. Such preorganization of the active site residues may be particularly important in the design of cysteine-rich metal-ion binding sites because of the proximity of physiological pH values to the pK_a of the free ligand.

Implications to the Stability of [4Fe−4S]$^{2+}$-IGA

The determination of the thermodynamic stability of the [4Fe−4S]$^{2+}$-**IGA** complex could not be performed by direct [4Fe−4S]$^{2+}$ titrations because of the instability of the iron−sulfur cluster in aqueous solution without supporting ligands, e.g., excess β-ME. However, the data presented for Fe^{II}-, Co^{II}-, and Zn^{II}-**IGA** and their correlation with the Irving−Williams series can be used to infer the practical limits of [4Fe−4S]$^{2+}$-**IGA** stability. The observation that the self-assembly reaction yields the [4Fe−4S]$^{2+}$-**IGA** complex rather than Fe^{II}-**IGA**, as evidenced by the spectrum in Figure 4C, is consistent with a greater thermodynamic stability of the former. This supposition is supported by the results of the proton competition experiment with [4Fe−4S]$^{2+}$-**IGA**. Titration of dilute acid into an anaerobic solution of [4Fe−4S]$^{2+}$-**IGA** results in loss of the S → Fe(II) charge-transfer transition at 380 nm consistent with cluster decomposition. Figure 11 shows the pH titration fit to a single protonation event with a pK_a^{eff} value of 6.8; thus, removal of the [4Fe−4S] cluster from **IGA** follows a mechanism distinct from the mononuclear complexes. The observation of a single pK_a^{eff} value is consistent with protonation of either a cysteinate or a μ_3-sulfido ligand followed by cluster hydrolysis. In either case, it appears that the first protonation

(161) Jiang, J.; Nadas, I. A.; Kim, M. A.; Franz, K. J. *Inorg. Chem.* **2005**, *44*, 9787−9794.
(162) Zagrovic, B.; Snow, C. D.; Khaliq, S.; Shirts, M. R.; Pande, V. S. *J. Mol. Biol.* **2002**, *323*, 153−164.
(163) Cerasoli, E.; Sharpe, B. K.; Woolfson, D. N. *J. Am. Chem. Soc.* **2005**, *127*, 15008−15009.

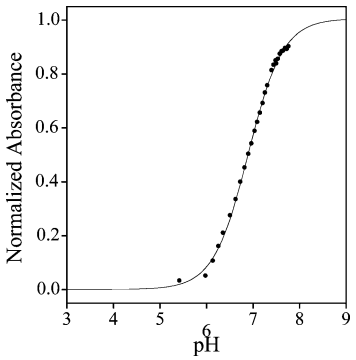


Figure 11. Spectroscopic determination of the competition between $[4Fe-4S]^{2+}$ and proton binding to **IGA**. Titration of acid into 54 μM $[4Fe-4S]^{2+}$-**IGA** results in loss of the absorption intensity as the pH is lowered. The data for the $[4Fe-4S]^{2+}$-**IGA** fit to a single protonation event with a pK_a^{eff} value of 6.8.

event of $[4Fe-4S]^{2+}$-**IGA** occurs at a pH value close to that of Co^{II}-**IGA** and pK_a^{eff} values of 6.4 and 6.7, respectively. This suggests that the binding of **IGA** to the $[4Fe-4S]^{2+}$ cluster is nearly as strong as the binding of **IGA** to Co(II). Because the stabilities of M^{II}-**IGA** follow the Irving−Williams series, we would propose that the $[4Fe-4S]^{2+}$ cluster might reside near Co(II) in this series.

Conclusions

Metallobiomolecules are highly elaborated coordination complexes,[164] and their fundamental metal−ligand interactions are critical components of metalloprotein folding, assembly, stability, electrochemistry, and catalytic function. While rigorous studies of metal−ligand coordination equilibria appear to have gone out of fashion,[34] this study clearly demonstrates their value in the bioinorganic chemistry of metalloprotein design, metal-induced protein folding, and metal-ion trafficking. Because designed peptide and protein ligands are employed as ever-more sophisticated synthetic analogues or maquettes of natural protein active sites, rigorous descriptions of their coordination chemistry are essential for accurate predictions of metal-ion binding-site affinities and selectivities in designed metalloproteins. Furthermore, the role of solution pH, relative to ligand pK_a values, in attenuating metal-ion affinity and selectivity is crucial for the precise definition of metal−peptide complex speciation in an aqueous solution required to fully comprehend the transport of metal ions in biological systems.

Herein, we have utilized traditional coordination chemistry methods[124] to define the metal-ion binding properties of an unstructured linear peptide containing four cysteines toward Fe(II), Co(II), Zn(II), and a $[4Fe-4S]^{2+}$ cluster. The data demonstrate that Fe(II), Co(II), and Zn(II) binding to the $(S-Cys)_4$ peptide provides −11.8, −15.8, and −21.6 kcal/mol of stabilization to the resulting metalloprotein structure, respectively. The femtomolar affinity of **IGA** for Zn(II) evidences that simple peptide ligands can have metal-ion formation constants higher than those of natural protein ligands. These affinities may reflect the maximal values possible for binding these metals to a tetrahedral tetracysteinate binding site. These metal binding energies are an order of magnitude larger than those of typical protein−protein interactions[62] and illustrate the effect metal-ion binding can have on protein folding, assembly, and stability.

The data presented provide both a basis for examining the minimal metal-ion binding requirements of a peptide ligand and a method to evaluate its thermodynamically preferred cofactor. Thermodynamic characterization of metal-ion affinities in protein ligands is a requisite first step in understanding the biosynthesis of natural metalloproteins. Only with a detailed understanding of the thermodynamics can one begin to reveal the kinetic factors involved in metal-ion uptake, release, and delivery that play crucial roles in the assembly of metalloproteins in vivo.[165] In other words, evidence for kinetic control of metalloprotein assembly in vivo relies on the observation of a deviation from the predictions based on metal−protein thermodynamics. The challenge is to elucidate both the thermodynamic energetics and the kinetic mechanisms that nature employs to orchestrate native metal-ion incorporation into metalloproteins.

Acknowledgment. A.K.P. and A.R.R. are National Science Foundation GK-12 Fellows (Grant DGE-02-31875). B.R.G. is a Camille Dreyfus Teacher−Scholar. This work was supported by the American Chemical Society Petroleum Research Fund (41550-AC3 to B.R.G. with a Summer Research Fellowship supplement for A.G.H.).

Supporting Information Available: The Material and Methods section and the derivation of the equations used to fit the various data sets. This material is available free of charge via the Internet at http://pubs.acs.org.

IC052190Q

(164) Ibers, J. A.; Holm, R. H. *Science* **1980**, *209*, 223−235.

(165) Ledwidge, R.; Soinski, R.; Miller, S. M. *J. Am. Chem. Soc.* **2005**, *127*, 10842−10843.

(166) Bau, R.; Rees, D. C.; Kurtz, D. M.; Scott, R. A.; Huang, H. S.; Adams, M. W. W.; Eidsness, M. K. *J. Biol. Inorg. Chem.* **1998**, *3*, 484−493.

(167) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 51−55.

(168) Banci, L.; Bertini, I.; Del Conte, R.; D'Onofrio, M.; Rosato, A. *Biochemistry* **2004**, *43*, 3396−3403.

(169) (a) Shi, W.; Dong, J.; Scott, R. A.; Ksenzenko, M. Y.; Rosen, B. P. *J. Biol. Chem.* **1996**, *271*, 9291−9297. (b) Wright, J. G.; Johnson, S. L.; Utschig, L. M.; O'Halloran, T. V. *J. Inorg. Biochem.* **1990**, *43*, 51. (c) Song, L.; Caguiat, J.; Li, Z.; Shokes, J.; Scott, R. A.; Olliff, L.; Summers, A. O. *J. Bacteriol.* **2004**, *186*, 1861−1868.