De novo designed metalloenzymes: structural stabilization and hydrolytic catalysis in a family of α-helical coiled coils

by

Melissa Lynn Zastrow

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in the University of Michigan
2013

Doctoral Committee:

Professor Vincent L. Pecoraro, Chair
Professor Hashim M. Al-Hashimi
Associate Professor Nicolai Lehnert
Professor Stephen W. Ragsdale

To my parents

Acknowledgements

There are many people without whom this thesis would not have been possible and who have wholly supported my graduate studies here at the University of Michigan. First, I would like to express my deepest gratitude to my primary research advisor, Professor Vincent Pecoraro, for his support, encouragement, guidance, advice, and especially his contagious enthusiasm for scientific research. Vince has provided me with a highly motivating atmosphere for doing research while encouraging independence. I am very grateful for all of the opportunities he has offered me, including a 10-week research sabbatical at Oxford University and my attendance at several Gordon conferences, both of which have certainly shaped my career. I consider myself extremely lucky to not only have been able to pursue the exciting project described in this dissertation, but to have had Vince as my research advisor, and I thank him for the significant role he has played in my life and career.

In addition to my thesis advisor, I would like to thank the rest of my thesis committee, Professors Nicolai Lehnert, Stephen Ragsdale, and Hashim Al-Hashimi for all of their hard questions, insightful comments, and enthusiasm for my research. I really appreciate all of the valuable time they have spent serving as my committee members.

I am indebted to Dr. Jeanne Stuckey and Dr. Jennifer Meagher who have taught me how to grow protein crystals, collect data, solve, and refine structures. This thesis would not be what it is without these crystal structures. I am very thankful for all of their assistance.

I would like to acknowledgement the NIH Chemistry-Biology Interface Training Program for providing me with a training grant and, more importantly, much invaluable exposure to biological research that has played a huge role in my research interests and fostered many valuable discussions with fellow trainees, especially at chalk talks and the yearly symposium.

I wish to thank Professor Fraser Armstrong and his lab members for welcoming me at Oxford University for my CBI sabbatical where I learned how to run protein film voltammetry experiments and for all of the memorable dinners, drinks at the pubs, and teaching me how to punt.

I owe a special thanks to Dr. Matteo Tegoni not only for the preliminary copper-peptide work he did which supported my research sabbatical proposal but also for his constant support, advice, and continuing friendship.

There are many Pecoraro lab members, both past and present, who I am happy to thank. First, I cannot thank Dr. Anna Peacock enough for being such a fun, enthusiastic (dancing over exciting data in the NMR room), and supportive mentor during my research rotation and encouraging me to join the Pecoraro lab. I am especially thankful to Dr. Saumen Chakraborty, Dr. Ted Boron, and Dr. Joe Jankolovits for all of the support, advice, and feedback they provided me with, for the many coffee breaks and happy hours, and for their continued friendships. I thank current group members Fangting, Leela, Jeff, Ginny, Evan, Cathy, Alison, Hira, and David for all of their support and encouragement throughout my time here. I thank Kathy for all of her help throughout my graduate studies. I thank any other members of the Pecoraro group or visitors with whom I have overlapped that I have not acknowledged by name.

I offer my thanks to all of my friends, in Michigan and elsewhere, for all of the wonderful times and memories we have shared over the years, from happy hours and dinners to coffee breaks and football games and tailgating. These experiences have truly enriched my time here.

Finally, I would like to thank my family for all of their love and support. In particular, I am extremely thankful to my parents for their continuous encouragement in my progress and for being so understanding and always willing to listen. This would not have been possible without them. I thank my brother, Derick, and my brother- and sisters-in-law, Eric, Marit, and Sara. I thank my parents-in-law for all of their love and support. Last but not least, I am extremely thankful to my loving husband, Gary, for all of his understanding and support, for his patience, for cooking way more than his fair share of dinners, and for being my constant anchor and a source of inspiration and motivation throughout these past five years.

Table of Contents

List of Tables

List of Figures

Abstract

Fundamental studies on structural stabilization and hydrolytic catalysis in a family of *de novo* designed α-helical three-stranded coiled coils are presented. Basic principles are used to discriminate structural and catalytic sites ($Hg(II)Cys_3$ and $Zn(II)His_3$). Spectroscopy, X-ray crystallography, and kinetic studies are exploited to characterize these sites.

X-ray crystal structures of these metallopeptides ($[Hg(II)]_S(Zn(II)(OH_2/OH^-)]_N$(**CS**L9PenL23H)$_3^{n+}$) are determined to 2.2 Å resolution (pH 7.5 and 8.5). Both structures contain a tetrahedral ZnN_3X site, like the active site of carbonic anhydrase (CA). At low pH a mixture of HgS_2 and HgS_3 centers is observed, while the high pH form exclusively shows HgS_3. Denaturation studies verify the HgS_3 site provides structural stability, while kinetic analysis of $[Hg(II)]_S[Zn(II)(OH_2/OH^-)]_N$ (**TRI**L9CL23H)$_3^{n+}$ towards non-physiological substrate, *p*-nitrophenyl acetate (*p*NPA), confirms the ZnN_3O site is hydrolytically active with a maximal efficiency (31 M^{-1} s^{-1}) only ~100-fold less than the fastest CA and pK_a ~2 units higher (9.0 vs 6.8). More significantly, analysis towards physiological substrate (CO_2) confirms the design of a minimal Zn(II)-binding site with a catalytic efficiency (~2.6 x 10^5 M^{-1} s^{-1}) within ~400-fold of a highly evolved enzyme. This minimal active site structure becomes even more competitive with CAII mutants that exclude important secondary hydrogen-bonding interactions.

Variation of the position of the ZnN_3O site along the sequence of the peptides (**TRI**L9HL23C, **TRI**L9CL19H) and insertion into a longer, more stable sequence (**Grand**L2WL16CL30H) demonstrates that while location can be used to control binding affinity, solvent, substrate, and inhibitor access, Zn(II)-OH_2 pK_a, and catalytic rate, the maximal catalytic efficiency remains significantly high. Kinetic analysis in the absence of the structural site ($[Zn(II)(OH_2/OH^-)]_N$(**TRI**L2WL23H)$_3^{n+}$) yields similar parameters demonstrating that this extra stability is not detrimental to the function of the second site.

Polar residues (Thr, Asp) are incorporated into **Grand** (**Grand**L16CL26TL30H, **Grand**L16CL26DL30H) for implementing secondary interactions involving hydrogen bonding in order to fine-tune the designed hydrolytic site. These substitutions do not result in large changes to the catalytic efficiency; however, Zn(II)-binding affinities increase ~70-fold. Co(II) binding (as a spectroscopic probe) to these variants indicates different and probably mixed coordination geometries relative to $[Co(II)(OH_2/OH^-)_x]_N$ (**TRI**L2WL23H)$_3^{n+}$, supporting the need for asymmetric site design in future work.

Chapter I. Designing functional metalloproteins: from structural to catalytic metal sites

Introduction

Protein design is a powerful approach for understanding metal sites in biology through reproduction of coordination environments and active site structures within a peptidic scaffold. Metal sites can efficiently catalyze many of the most complex and essential reactions in nature.[1,2] The ability to reproduce these reactivities in a designed system is a true test of our knowledge of protein folding and how structure and function are related in metalloproteins. Relative to protein design in general, designing metal sites into proteins can be challenging due to the variability of metal ions in their oxidation states, ligand donor sets, and preferred geometries. However, many metal sites can also serve as spectroscopic tags to assist in characterization and assessment of the success of the design, thus speeding up the design process. Additionally, the variability and sheer catalytic power inherent in metal ions increases the variety of functions that can be achieved, which may allow for a wider range of novel metalloenzymes for commercial use. Since the first reports of designed proteins and metalloproteins began appearing over 30 years ago, significant progress has been made in reproducing many native metal sites for a large variety of metals and metal cofactors.[3–11] However, designed metalloproteins have often operated with a significant lack of catalytic efficiency relative to their designed counterparts.[6,8,12] Now, there are several examples of functional metalloproteins with increasing enzymatic activities, suggesting that rivaling native metalloenzymes is a more accessible goal than ever. Reaching this point, however, has required many advances in metalloprotein (and non-metalloprotein) folding and design. Specifically, one must be able to design structural sites where a metal can impart thermodynamic stability, direct protein folding and/or stabilize non-preferred metal-binding geometries. These advances in understanding the balance between the protein scaffold stability and metal

ion coordination preferences often naturally lead to the design of catalytically active metal sites using a variety of approaches. The ultimate goal is to generate efficient metalloenzymes through a bottom up (*de novo* or "from scratch") approach in which a primary sequence is chosen that folds into the predicted structure (requiring a deep understanding of protein folding, structure, and function). Other strategies, such as the redesign of native protein scaffolds, also provide new insight for metal-protein interactions. Through these complementary approaches, one might expect to be able to apply the insights gained throughout these studies to achieve one of the larger goals of the field: the design of novel metalloproteins with improved properties (stability, efficiency) and/or new catalytic activities. Such constructs, which are not found in nature, may be useful for a range of applications, including industrial processes where they could represent cheaper, more stable, and environmentally friendly alternatives to the current catalysts.

Metals in proteins have a wide range of functions that can be broadly classified into a few types: structural, catalytic, and electron transfer. Understanding how the protein matrix can fine-tune the properties of metal ions to affect such a variety of functions is an important question being actively addressed by protein design. The design of metalloproteins with electron transfer capabilities has been described.[4,13] Here, I will focus on designed metalloproteins as they relate to structural and catalytic metal functions. After a discussion of general protein design approaches and scaffolds, I will described the design of structural metal sites (Section 2), highlighting the Pecoraro group's work on the *de novo* design of metal-binding three-stranded coiled coils (3SCCs) (Section 2b), and discuss examples of functional *de novo* designed metalloenzymes (Sections 3 and 4). Section 3 will describe examples of designed enzymes which are not Zn(II)-dependent. Section 4 will describe the design of Zn(II)-binding proteins towards the generation of efficient hydrolytic zinc enzymes.

1. Metalloprotein design approaches and scaffold choices

Based on selection of the starting scaffold, metalloprotein design approaches can be categorized broadly into protein redesign or *de novo* ("from scratch" in its purest

form) protein design. Metalloprotein redesign encompasses a variety of approaches including redesign using non-metalloprotein scaffolds as a starting point and using existing metal sites and redesigning them into new metal sites with altered properties. *De novo* metalloprotein design, on the other hand, requires both design of a unique protein scaffold and incorporation of the metal site(s). Computational and/or empirical approaches to the design can be taken. Computational design of metalloproteins is currently limited due to the inherent variability in metal sites and a lack of suitable parameters to describe them (especially for the design of metal sites with non-preferred geometries), but nonetheless there are successes in this area.[5,7,14,15] Empirical design can be less limiting but requires more knowledge about the templates being used and the design target. A combination of the two is often used. Further, proteins and metalloproteins can be designed combinatorially or rationally. Combinatorial design is typically used when little is known about the design goal and involves preparing a library of designs to sample, although it can be difficult to sample enough sequence variety.[5] Rational design is the ideal approach and allows more control over the outcome but also requires more in-depth knowledge of the structure and function. This approach has seen much success with the design of heme proteins.[3] Often, a combination of rational and combinatorial approaches can be a very effective strategy. Here, I will focus on rationally designed metalloproteins.

a) Protein redesign

One can take a redesign approach by beginning with a native protein scaffold and incorporating a novel metal site, through introduction of the necessary ligands into a non-metalloprotein (or into a new location of an existing metalloprotein), redesign of an existing metal site, or incorporation of a metal cofactor or complex. Evolved proteins tend to be fairly stable and can handle the multiple mutations required to engineer a metal site. If chosen correctly, they can be easily obtained in high yield and are already well characterized. It can even be argued that nature uses this approach. Although there are a large number of characterized sequences and structures available, there are only ~1000 unique protein folds known and often, many different types of proteins with diverse

functions use a single protein fold (an important example of this is the Greek key β-barrel fold).[5] It is possible that nature may only use a small number of thermodynamically stable scaffolds to design repeatedly many proteins with diverse active site structures and functions. Relative to *de novo* design, this approach can currently offer a wider variety of potential scaffolds, as there are relatively few robust *de novo* designed scaffolds (Section 1c). On the other hand, this could also prove to be a drawback as it may be difficult to divulge hidden structural features in complex native proteins. Nevertheless, there is much that has been and will continue to be learned from the redesign of existing metalloproteins. Although my focus is on *de novo* designed proteins, there are several recent and impressive reports of redesigned functional metalloproteins[14,16–25] and many others that have been reviewed.[3,5,7,9,26–30]

b) De novo protein design

The *de novo*, or in its purest form, "from scratch", approach to metalloprotein design is to build a well-folded protein structure from a primary sequence not directly related to that of any natural protein and subsequently engineer a metal-binding site.[10] The basis of *de novo* design usually involves a minimization or simplification of protein structure.[31,32] One of the goals is to attain a minimal specific structure that removes much of the excess of a native protein but retains sufficient complexity for protein-like functions. The main advantage of this approach is that it eliminates most of the convoluted behaviors that make characterizing many native metalloproteins difficult. It is likely that this simplified approach to characterizing metal binding and function may uncover otherwise hidden subtleties in structure-function relationships. With a deeper understanding of the necessary features, increasing amounts of complexity can then be incorporated, as deemed necessary, into these minimal designs. Ultimately, careful analyses of *de novo* designed metalloproteins, combined with the information learned from redesign endeavors, should lead to the ability to prepare *de novo* designed metal-based biocatalysts for a variety of desired functions not necessarily found in nature and for a variety of applications.

c) Scaffolds for designing metal-binding sites

Design of a metal-binding site, regardless of the approach, first requires a stable protein scaffold into which mutations can be made to incorporate ligands for metal (or metal cofactor) binding. For protein redesign, the scaffold is typically chosen based on the goal of the design and the properties of the native scaffold and nature provides many existing scaffolds to choose from.[5] A variety of approaches can be taken including engineering a metal- or metal cofactor-binding site into a non-metalloprotein[3,7,14,16,17,26–28], redesigning an existing metal-binding site[19,33–35], removing and inserting entire loop sequences containing the desired metal-binding site[5,9,20], and generating multimeric protein structures with metal-mediated protein interfaces.[18,36–39]

Far fewer *de novo* designed scaffold choices have been prepared to date, which limits the range of metalloprotein environments that may be examined. Ultimately, one goal of *de novo* design would be to expand this repertoire to include at least as many, if not more, scaffolds than those provided by nature. Initial attempts at *de novo* protein design lead to structures that formed molten globules, but now there are a number of successful examples of small uniquely folded proteins. Hydrophobicity provides a strong driving force for folding but many more subtle features also need to be incorporated for both stability and specificity such as hydrogen bonding, electrostatics, side-chain

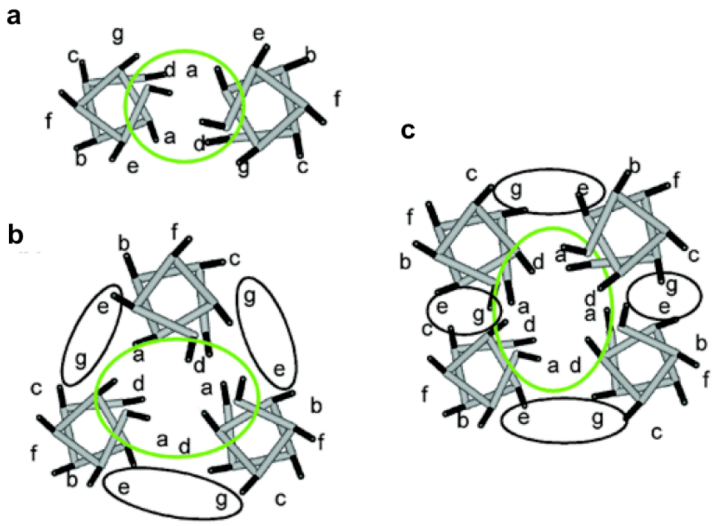

Figure 1-1. Helical wheel diagrams for parallel (a) two-, (b) three-, and (c) four-stranded coiled coils. Figure reproduced from ref. 53.

packing, and conformational preferences. Most *de novo* designed proteins consist of α-helices and are either coiled coil structures made up of self-assembling peptides or α-helical bundles made up of helix-loop-helix motifs.[8,10,40] The fact that α-helices have intrastrand interactions with residues that are close in the primary sequence to stabilize their structures is one aspect (as opposed to interstrand hydrogen bonding found in β-sheets) that made them an attractive initial target for design.

Hodges and coworkers pioneered the *de novo* design of α-helical coiled coils.[10,41–44] The α-helical coiled coil is made up of a bundle of parallel or antiparallel amphipathic α-helices with a left-handed superhelical twist.[45–48] The primary sequence contains a seven residue repeating sequence (the heptad repeat), with positions designated **a**-**g**. The residues in the **a** and **d** positions are directed towards the interior of the structure and tend to be hydrophobic while hydrophilic residues in the remaining positions are involved in electrostatic interactions. The association of hydrophobic **a** and **d** residues provides part of the driving force for assembly of the coiled coil. Helix-inducing residues, such as Ala, may be placed in the **c** position. Figure 1-1 illustrates these interactions for parallel two-, three-, and four-stranded coiled coils (XSCCs where X = 2, 3, or 4). Specific oligomers of α-helices can be attained through the rational design of core-packing interactions (typically by varying the size of the hydrophobic residues) and surface electrostatics (which also contribute to stabilization of parallel or antiparallel configurations of coiled coils).[32,42,49–52] The heptad repeat approach is a highly versatile method that has yielded many successful designed metalloproteins (Figure 1-2a, b).[7,53–70] In particular, the Pecoraro group has designed the **TRI** family of peptides, which are α-helical coiled coils that fold into specific aggregates depending on pH, peptide length, and sequence composition (these will be described in detail in Section 2b).[53,63,71,72]

Another structure type extensively explored in protein design is the α-helical bundle. Although often designed based on coiled coils, α-helical bundles can be structurally different in that the helices gradually diverge from a point near one end of the bundle.[73,74] The most common are the four-helix bundles and one of the earliest examples of a designed four-helix bundle is α_4.[31,75] This small protein was prepared by taking short helical sequences that could self-associate into a tetramer, adding hairpin loops to make a dimer of helix-loop-helix motifs, and finally incorporating a third loop to create a single

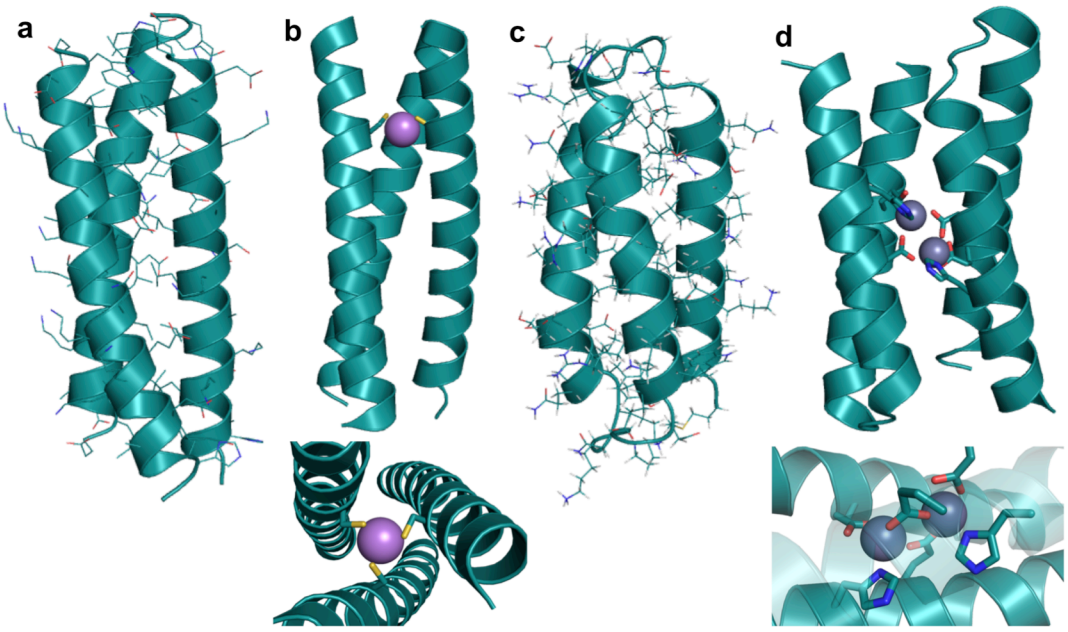

Figure 1-2. Structures of representative *de novo* designed α-helical coiled coils and bundles. (a) X-ray structure of a 3SCC, CoilSer, generated from pdb 1COS[50], (b) X-ray structure of a metal-bound 3SCC, $[As(III)]_3(CSL9C)_3$, generated from pdb 2JGO[68] with side-on view (top) and top-down view (bottom); As-S metal-ligand bond distances all 2.3 Å, (c) NMR structure of a three-helix α-helical bundle protein generated from pdb 2A3D[74], (d) X-ray structure of a dinuclear metal site in a designed four-helix bundle protein (diZn(II)-DF1) generated from pdb 1EC5[15] with side-on full view (top) and close-up view of metal site (bottom); All metal-ligand bond distances are in the range 1.8-2.1 Å.

primary sequence. Although this bundle adopted a folded, globular conformation in aqueous solution, the core was too conformationally flexible (it contained only Leu residues) leading it to exhibit some molten globule-like characteristics rather than assuming a single native state. Insertion of metal-binding residues to create a $ZnHis_3$ site did help induce a more native-like state into α_4.[49,76] Now, there are many examples of well-folded α-helical bundles that adopt a single native state due to the incorporation of a variety of more specific interactions both in the hydrophobic core and on the hydrophilic surface.[10,74,77] There are also examples where metal sites are engineered into α-helical bundles.[15,76,78–80] The Pecoraro group has recently reported the incorporation of a heavy metal-binding Cys_3 site into a three-helix α-helical bundle (α_3D, Figure 1-2c) that was originally designed based on **CoilSer** (**CS**, a *de novo* designed 3SCC, Table 1-1, Figure 1-2a, b).[74,79] DeGrado and coworkers have reported an extensive amount of work on the

dimetal-containing four-helix bundle Due Ferri (DF, Figure 1-2d, to be discussed in detail in Section 3a).[6,15,81–84]

Although the *de novo* design of proteins has been dominated by preparation of α-helical coiled coils and bundles, there has also been substantial effort directed towards the formation of β-sheet structures. This is relatively challenging, not only because of the need to create more interstrand interactions between residues well-separated in sequence, but because there are issues with aggregation, overall stability, folding cooperativity, and challenges in designing turn sequences. Residues that are bulky and hydrophobic, such as Tyr, Phe, Thr, Ile, Val, and Trp, seem to have a higher propensity towards the formation of β-sheet structures. Despite these challenges, there are several successes in the design of β-sheet structures[85–90] and related scaffolds including mixed α/β structures[91–94], γ turns[95], β-hairpin peptides[96–99], and small peptides with loop region metal clusters.[100,101] There is also limited success in the *de novo* design of β-structured scaffolds for functional metalloproteins; the design of a redox-active rubredoxin mimic is an example of one of these rare β-sheet metalloproteins.[102]

Table 1-1. CoilSer (CS) and TRI peptide family sequences used in these studies.

Peptide	a b c d e f g	
CoilSer (CS)	Ac-E WEALEKK LAALESK LQALEKK LEALEHG	-NH_2
CSL9C	Ac-E WEALEKK **C**AALESK LQALEKK LEALEHG	-NH_2
CSL19C	Ac-E WEALEKK LAALESK LQA**C**EKK LEALEHG	-NH_2
CSL9PenL23H	Ac-E WEALEKK **Pen**AALESK LQALEKK **H**EALEHG	-NH_2
TRI	Ac-G LKALEEK LKALEEK LKALEEK LKALEEK	G-NH_2
TRIL9C	Ac-G LKALEEK **C**KALEEK LKALEEK LKALEEK	G-NH_2
TRIL12C	Ac-G LKALEEK LKA**C**EEK LKALEEK LKALEEK	G-NH_2
TRIL16C	Ac-G LKALEEK LKALEEK **C**KALEEK LKALEEK	G-NH_2
TRIL12CL16C	Ac-G LKALEEK LKA**C**EEK **C**KALEEK LKALEEK	G-NH_2
TRIL2WL9C	Ac-G **W**KALEEK **C**KALEEK LKALEEK LKALEEK	G-NH_2
TRIL2WL23C	Ac-G **W**KALEEK LKALEEK LKALEEK **C**KALEEK	G-NH_2
TRIL23H	Ac-G LKALEEK LKALEEK LKALEEK **H**KALEEK	G-NH_2
TRIL2WL23H	Ac-G **W**KALEEK LKALEEK LKALEEK **H**KALEEK	G-NH_2
TRIL9CL23H	Ac-G LKALEEK **C**KALEEK LKALEEK **H**KALEEK	G-NH_2
Baby	Ac-G LKALEEK LKALEEK LKALEEK	G-NH_2
BabyL9C	Ac-G LKALEEK **C**KALEEK LKALEEK	G-NH_2
BabyL12C	Ac-G LKALEEK LKA**C**EEK LKALEEK	G-NH_2
Grand	Ac-G LKALEEK LKALEEK LKALEEK $(LKALEEK)_2$	G-NH_2

N- and C-termini are capped by Ac and NH_2 groups (acetylated and amidated), respectively.

2. Designing structural metal sites

Many metal ions play a structural role in their protein environments. They can act to simply stabilize a structure by inducing or assisting folding or as switches for converting one conformation into another (for example, to regulate activity). These are most commonly represented by Zn(II)-thiolate sites (as in zinc fingers which are typically comprised of $ZnCys_2His_2$ coordination environments, and will be discussed in more detail in Section 4) and calcium sites (calmodulin), although others exist.[103–105] There are examples of designed metal-binding sites in which the metal plays a structural role.[53,64,70] In addition to describing the Pecoraro group's work involving heavy metal binding, particularly with Hg(II) to thiol sites in α-helical coiled coils in Section 2b, I will briefly highlight other examples and recent approaches to the design of structural metal-binding sites in proteins.

a) Metal-assisted and metal-induced folding in α-helices and α-helical coiled coils

In the field of metalloprotein design, there are many examples in which metal ion-assisted folding and self-assembly of α-helical coiled coils generates well-defined secondary and tertiary structures.[64,69,106–111] Although these studies are extensive, here I will focus on those which are relevant to the most recent studies in the context of this chapter. Further discussion on those designs which are Zn(II)-dependent will be provided in Section 4a. Figure 1-3 provides models of some of the important constructs described below.

Ghadiri *et al.* were among the first to exploit the design of metal-binding sites in peptides to promote the formation of α-helices[112,113] and coiled coils.[107,114] They designed a 17-mer peptide (Ac-AEAAAKEAAAKX_1AAAX_2A-$CONH_2$ where X_1 = His and X_2 = His or Cys) which could bind Cd(II), Cu(II), Ni(II), and Zn(II), all with an enhancement in α-helicity from 54% to ~90%. Using the His/His version of this peptide, binding of an exchange-inert metal complex, $[Ru(NH_3)_5(OH_2)]^{2+}$, leads to the formation of a macrocyclic complex, *cis*-$[Ru(NH_3)_4(His)_2peptide]^{3+}$, with 80% α-helical content. The

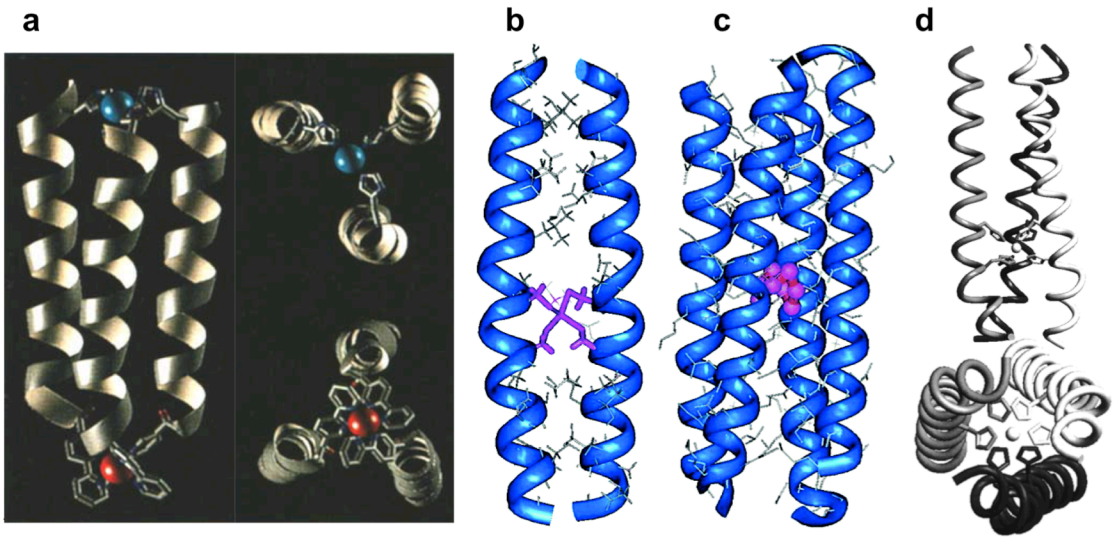

Figure 1-3. Examples of metal-stabilized α-helical structures described in this section. (a) Computer-generated model of the parallel three-helix bundle Cu(II)Ru(II) metalloprotein with full side-on view (left) and close-up views of each metal site (right). Figure reproduced from ref. 114. (b) Energy-minimized computer model of the Cd(II)-bridged C16C19 peptide dimer (2SCC) with Cd(II) bound in a tetrahedral geometry. Figure reproduced from ref. 13. (c) Computer-generated model of the tetrameric Cu(I)-C16C19-GGY metalloprotein (4SCC). Figure reproduced from ref. 13. (d) Model of the Ni(II)-His6 complex of IZ-3adH (3SCC) with full side-on view (top) and a bottom-up view from the C-termini (bottom). Figure reproduced from ref. 110.

group has also designed a 15-mer amphiphilic peptide with a 2,2'-bipyridine moiety at the N-terminus which, upon binding Ni(II), Co(II), or Ru(II), self-assembles to form a 3SCC with > 70% α-helical content.[107] This construct was also used as a template for incorporation of a Cu(II)$(His)_3$ site to form a heterodinuclear three-helix construct (Figure 1-3a).[114] Further, a stable parallel four-helix bundle metalloprotein is obtained by incorporating another 15-residue amphiphilic peptide with a pyridyl group at the N-terminus and binding Ru(II).[115]

Along similar lines, the Ogawa group has also reported the use of metal-binding sites to generate α-helical bundles. Kharenko *et al.* designed a 32-mer peptide C16C19-GGY (Ac-K(IEALEGK)$_2$(CEACEGK)(IEALEGK)GGY-$CONH_2$), which has Cys residues in adjacent **a** and **d** positions of one of the heptads, resembling the Cys-X-X-Cys motif found in rubredoxin. Upon binding Cd(II), the apo-peptide went from a random coil to a 2SCC containing a single Cd(II) ion (Figure 1-3b).[13,54] Interestingly, when Cu(I) is bound to the same peptide, a four-helix bundle metalloprotein containing a tetranuclear Cu_4S_4 cofactor is generated (Figure 1-3c).[116] Further, upon binding Pt(en)$(NO_3)_2$ to the

30-residue peptide AQ-Pal14Pal21 (which contains two 4-pyridylalanine, or Pal, binding residues on its surface and exists as a 2SCC in its apo form) results in a significant conformational change to a metal-bridged four-helix bundle.[108]

Tanaka and coworkers have used the heptad repeat approach to design a parallel 3SCC peptide (in the presence of bound metal ions) with the sequence YGG(IEKKIEA)$_4$ called IZ. The peptide IZ-3aH was generated by substitution of an isoleucine residue on the third heptad with a His and IZ-3adH could be generated with two His substitutions.[110,117] While this sequence showed only random coil-like secondary structure in solution, it was induced to fold into a 3SCC in the presence of metal ions such as Cu(II), Zn(II), Co(II), and Ni(II) (Figure 1-3d). Metal binding could be followed by monitoring the folding with circular dichroism (CD) as a function of the metal concentration, yielding dissociation constants in the low μM range. Substituting Cys residues into the IZ peptides also yields metal-binding sites for Cd(II) and Hg(II).[111] Recently, the group reported a study in which they fused the IZ-3adH peptide to another functional protein (a derivative of a naturally occurring DNA-binding domain) to create a metal ion-controlled DNA-binding protein which could modulate hydrolysis of the DNA site and inhibit RNA transcription by T7 RNA polymerase.[118]

As will be discussed in Section 2b, the Pecoraro group has shown that heavy metal-thiolate binding can also be used to induce and mediate the folding of 2- and 3SCCs.

Further, as will be discussed in Section 3b, the Ball group has exploited the metal-induced α-helicity of 2SCCs and the inherent properties of coiled coil structures to design dirhodium-based catalysts for a variety of roles relevant to enzymatic function.

b) Binding of Hg(II) to the TRI family of α-helical coiled coils

The Pecoraro group's work is based on the **TRI** family of *de novo* designed peptides[71,72] for which the parent sequence, Ac-G(LKALEEK)$_n$G-NH$_2$, is based on that of **CoilSer**[50,119] and consists of four heptad repeats. This heptad repeat sequence results in amphipathic α-helices with all Leu residues in the **a** and **d** positions on the hydrophobic face, Ala in the **c** position, and Lys and Glu residues in most of the remaining positions

on the hydrophilic face. In aqueous solution, the hydrophobic residues aggregate at low pH to form a 2SCC and as the pH is raised above 5.5, salt bridge interactions between Lys and Glu at the interfaces between helices stabilize a parallel 3SCC. The **a** and **d** residues point towards the interior of the coiled coil to form the hydrophobic core. Substitution of these Leu residues with residues such as Cys or His yields metal-binding sites. Figures 1-4 and 1-5 display the crystal structures of apo-(**CS**L9C)$_3^{n-}$ (**a** site) and apo-(**CS**L19C)$_3^{n-}$ (**d** site), respectively.[67] Peptide sequences are given in Table 1-1. These structures demonstrate the differences between these types of sites in the absence of metal. For the **a** site, the Cys side chains show alternate conformations with the predominant form preorganized for metal binding. Comparison of this apo-peptide with the corresponding metal-bound version, [As(III)]$_S$(**CS**L9C)$_3$[68] ([As(III)]$_S$ is used to

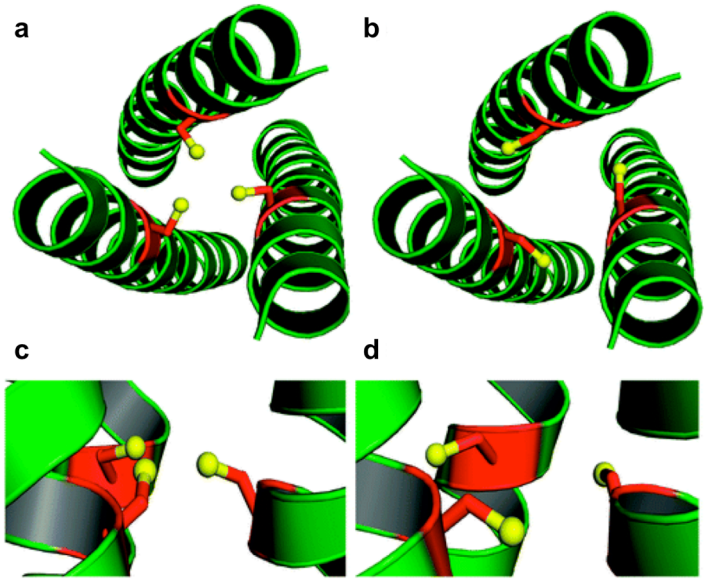

Figure 1-4. Ribbon diagrams of the X-ray crystal structure of apo-(CSL9C)$_3^{n-}$ showing the orientation of the Cys ligands. The Cys side chains are shown as red sticks with thiol groups colored yellow. Top-down views from the N-termini show (a) the orientation of the major conformer with all Cys side chains pointing towards the interior of the trimer and (b) the minor conformer with Cys side chains pointing towards the helical interface. Side-on views further demonstrate the flexibility of this site. (c) In the major conformer, the thiol groups point towards the N-termini and (d) in the minor conformer, the thiol groups point towards the C-termini. Figure reproduced from ref. 67.

indicate that As(III) is bound at the Cys_3 site), suggests only minor changes in the Cys residue geometries where they are slightly more constrained with metal binding. The **d** site residues, on the other hand, exist in a single conformation different than those observed in **CS**L9C, resulting in a larger cavity which can accommodate metals with larger ionic radii, such as Pb(II).

The Pecoraro group has also shown that non-natural amino acids can be used for metal recognition. Specifically, Penicillamine (Pen) can be used to increase steric bulk directly around the metal-binding pocket, thus enforcing lower coordination geometries than may otherwise be observed.[66,120,121] Further, alternate chirality amino acids have been shown to create smaller metal-binding pockets when placed in the Leu layer above the metal-binding site.[122] Moreover, placement of a D-amino acid ligand in the metal-binding Leu layer alters the ligand environment.[123] Specifically, Peacock *et al.* reported the crystal structures of apo-$(\mathbf{CS}L16Pen)_3^{n-}$ and apo-$(\mathbf{CS}L16DPen)_3^{n-}$, which demonstrate that incorporation of a D-amino acid ligand results in a metal binding site that is not preorganized for metal binding (Figure 1-6). This work provides a structural explanation for the relatively poor binding of Cd(II) to D-Pen as compared to L-Pen.

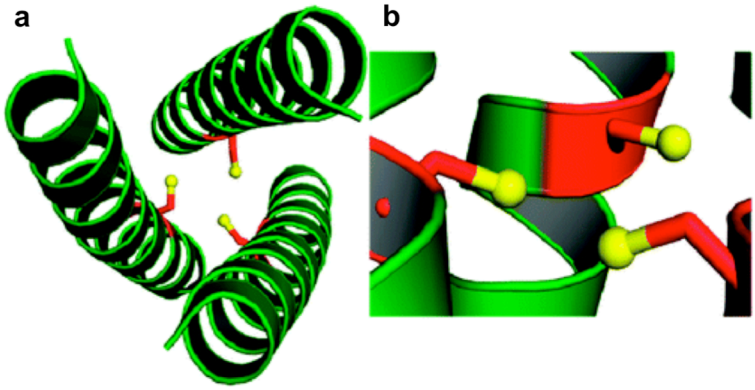

Figure 1-5. Ribbon diagrams of the X-ray crystal structure of apo-$(CSL19C)_3^{n-}$ showing the orientation of the Cys ligands. The Cys side chains are shown as red sticks with thiol groups colored yellow. (a) The top-down view from the N-termini shows that two of the Cys ligands point towards the interior of the coiled coil and the third toward the helical interface. (b) The side-on view demonstrates that two thiol groups point towards the C-termini while the third is almost perpendicular to the helical axis. Figure reproduced from ref. 67.

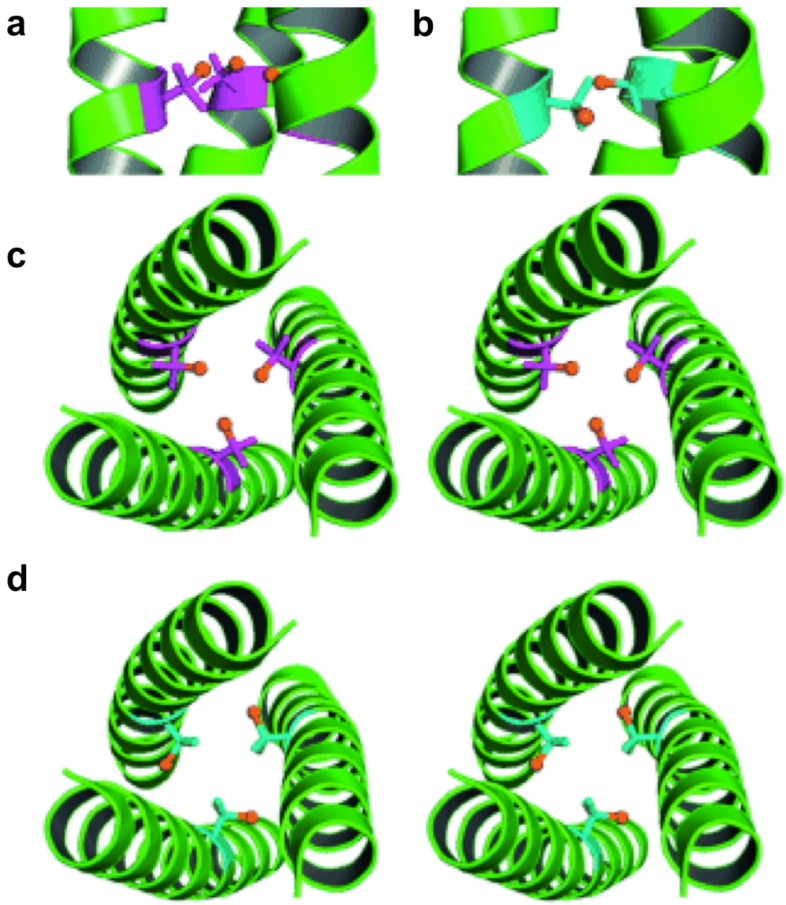

Figure 1-6. Ribbon diagrams of the X-ray crystal structures of apo-(CSL16Pen)$_3^{n-}$ and apo-(CSL16DPen)$_3^{n-}$ displaying the orientation of the Pen side chain. The L- (purple) and D-Pen (blue) side chains are shown in stick form with the thiol group colored orange. Side-on views show (a) L-Pen residues oriented toward the N-termini and (b) D-Pen residues oriented toward the C-termini. Top-down stereo views from the N-termini are shown and display (c) L-Pen residues oriented toward the interior of the coiled coil and (b) D-Pen residues oriented toward the helical interface. Figure reproduced from ref. 123.

In theory, the three-fold symmetric nature of the coiled coils can allow for the preparation of most metal geometries within its scaffold (Figure 1-7). With substitution of a single Leu layer, trigonal planar or trigonal pyramidal structures can be achieved depending on the stereochemical preference of the metal. A tetrahedral geometry may be observed with the inclusion of a solvent molecule and with two solvent molecules, a trigonal bipyramidal metal site might be observed. Substituting two adjacent Leu layers and inserting the appropriate metal may yield an octahedral site (Suzuki *et al.* have

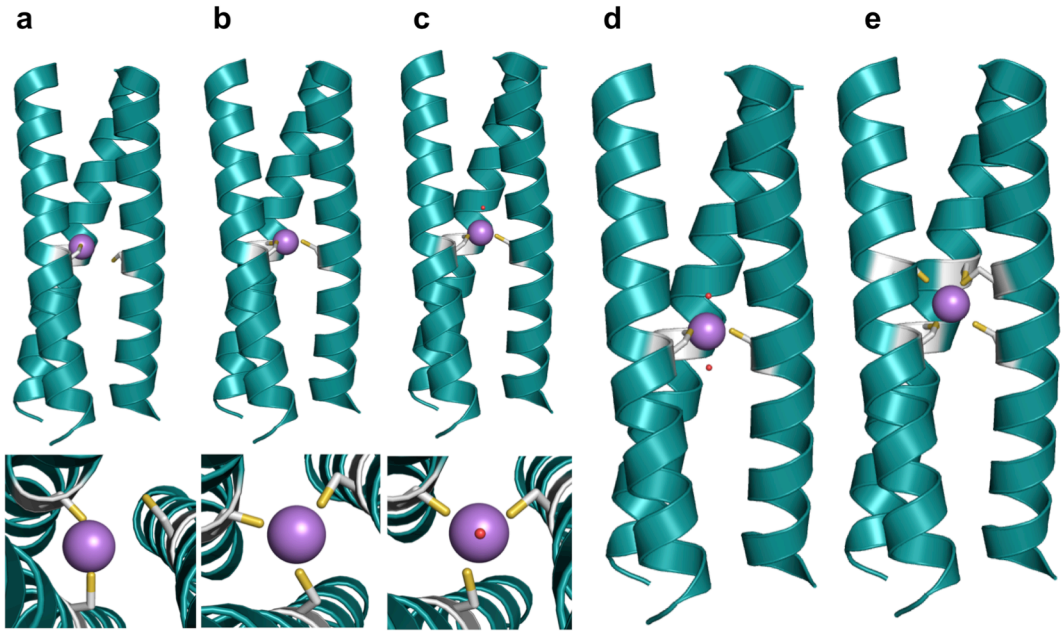

Figure 1-7. PyMOL[264] models of possible metal-binding geometries in a 3SCC. The metal-binding ligands are Cys residues and the metal is colored purple. (a) side-on view (top) and top-down view (bottom) of two-coordinate or linear geometry, (b) side-on view (top) and top-down view (bottom) of three-coordinate trigonal planar geometry, c) side-on view (top) and top-down view (bottom) of four-coordinate tetrahedral geometry, (d) side-on view of five-coordinate trigonal bipyramidal geometry, and (e) side-on view of six-coordinate octahedral metal geometry. Models are generated using the crystal structure of $[As(III)]_S(CSL9C)_3$ (pdb 2JGO)[68].

reported an α-helical trimer with two adjacent His_3 sites which can bind Ni(II), Zn(II), and Co(II) in octahedral geometries[110]). Further, a linear metal-binding site may be attained by either lowering the pH to stabilize the 2SCC, incorporating the appropriate metal and metal:peptide ratios (as will be described for Hg(II)), or at lower pH's in the presence of a 3SCC (where the third protein ligand present is not bound), as demonstrated in the structure of $[Hg(II)]_S[Zn(II)(OH_2/OH^-)]_N(\mathbf{CS}L9PenL23H)_3^{n+}$, which I will describe in Chapter 2 (the nomenclature of this complex now accounts for a coordinated solvent molecule to the Zn(II) ion: if the pH conditions are such that the solvent may occur as both H_2O and OH^- then OH_2/OH^- will be used). Therefore, using a three-fold symmetric coiled coil as a scaffold should allow for an extensive range of different metal-binding structures for a variety of metals and amino acid ligands within the same relative protein environment.

Until recently, most of the Pecoraro group's work on this system has focused on the preparation of thiolato coordination environments for metals such as Hg(II), Cd(II), Pb(II), As(III) and Bi(III) with the goal of understanding how heavy metals are coordinated in biological environments.[53,59–62,67,68,71,72,120,121,124–129] Although the toxic effects of these metals are well known, the causes are often not well understood due to a lack of suitable models, especially in the case of Hg(II).[63] Preparing thiophilic metal-binding sites in the **TRI** family of proteins has yielded not only a better understanding of how heavy metals are coordinated in native metalloproteins, but also information on how to selectively control these coordination environments. Previous reviews have discussed Hg(II) and Cd(II) binding in some detail[53,59,63,65], but here I will focus on describing Hg(II) binding as a means to achieve a structurally stabilizing metal-binding site in a 3SCC.

The Pecoraro group's protein design studies began by examining Hg(II) binding to peptides such as **TRI**L16C and **TRI**L12C, where Leu has been substituted for Cys in

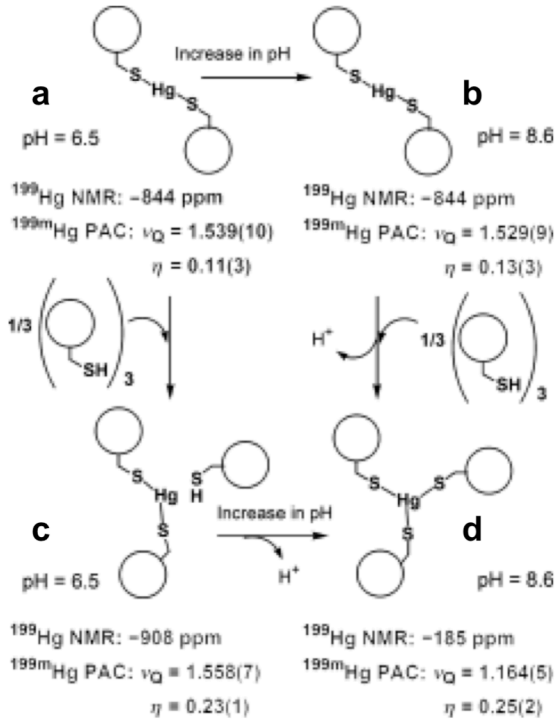

Figure 1-8. Species present at different TRIL9C/Hg(II) ratios and pH values. Figure reproduced from ref. 265.

an **a** site and a **d** site, respectively (sequences of the peptides to be discussed can be found in Table 1-1). As the group has shown with crystal structures of apo-(**CS**L9C)$_3^{n-}$ and apo-(**CS**L19C)$_3^{n-}$ (Figures 1-4 and 1-5), and, as will be described in terms of solution studies below, this seemingly slight shift in sequence position is significant for metal binding. To describe the binding of Hg(II) to these peptides, a variety of conditions must be considered (Figure 1-8). At low pH (< 5.5), 2SCCs containing a linearly-bound Hg(II) ion are formed. It is well known that this is the preferred geometry of Hg(II).[130,131] More interestingly, three distinct behaviors are observed at higher pH. When the peptide monomer (single α-helix) to Hg(II) ratio is 2:1, 2SCCs with two-coordinate Hg(II) are formed for both peptides, indicating that the Hg(II)-binding preference dominates the preferred bundle aggregation state. When this ratio is increased to 3:1, and in the pH range of 5.5-7, both peptides contain two-coordinate Hg(II) within a 3SCC, simply suggesting that the third Cys remains protonated and unbound. Finally, for **TRI**L16C above pH 8.6 and **TRI**L12C above pH 9.5, fully three-coordinate trigonal planar Hg(II) is obtained. Based on an equilibrium between two- and three-coordinate Hg(II) within the 3SCC, the pK_a for deprotonation of the third Cys can be determined and is 7.8 for the **a** site and 8.5 for the **d** site, indicating that metal binding to Cys residues is dependent on the **a** vs **d** substitution pattern. Importantly, this represents the first peptidic system to bind Hg(II) as a trigonal thiolato complex in aqueous solution and is an excellent model for the regulatory protein, MerR.[71,72,132] I have also reported the crystal structure of the complex (to be described in detail in Chapter 2), $[Hg(II)]_S[Zn(II)(OH_2/OH^-)]_N$ (**CS**L9PenL23H)$_3^{n+}$ (pH 8.5), which contains a trigonal HgS_3 site in one trimer of the asymmetric unit (ASU) and a T-shaped Hg(II) site in the other[70], crystallographically confirming the spectroscopic data. Through preparation of dual-cysteine peptides (in which two adjacent layers of Leu are substituted), the Pecoraro group was able to generate exclusively four-coordinate HgS_4 with the sequence **TRI**L12CL16C, representing a model of Hg(II)-substituted rubredoxin.[133,134] Extensive spectroscopic analysis of this designed system using numerous techniques for probing the metal coordination over a broad range of timescales (including UV-Visible characterization, ^{199}Hg NMR, ^{199m}Hg perturbed angular correlation (PAC, where ^{199m}Hg is a metastable isotope of ^{199}Hg which emits two successive γ rays upon its decay) and X-ray Absorption

(XAS) spectroscopies) provides a summary of the properties expected for homoleptic Hg(II)-thiolates in biological systems (Table 1-2).[63]

Table 1-2. Spectroscopic values for $[Hg(II)]_S$(TRILXC)$_X^{n-}$ complexes.

Hg coordination mode	λ (Δε)[a]	pK_a	R_{HgS} (Å)[b]	δ (^{199}Hg ppm)	PAC (ν_Q, η)[c]
Linear 2-coordinate	240 (2700)[d]	-	2.32[d]	-844[e]	1.53, 0.13[e]
Trigonal 3-coordinate **a** site	247 (19200)[f] 265 (11900)[f] 295 (5800)[f]	7.6 ± 0.2[g]	2.44[f], 2.23[l]	-185[e]	1.16, 0.25[e]
Trigonal 3-coordinate **d** site	230 (21300)[h] 247 (15000)[h] 297 (5500)[h]	8.5 ± 0.2[h]	2.44[h]	-316[i]	-
Linear 2-coordinate within a 3SCC	247 (2000)[j]	-	2.29, 2.13, 3.06[m]	-908[e]	1.56, 0.23[e]
Tetrahedral 4-coordinate	230 (8100)[k] 289 (7100)[k]	-	-	-500[k]	-

a. λ given in nm and Δε given as M^{-1} cm^{-1}.
b. Average Hg-S EXAFS bond lengths.
c. ν_Q given in GHz and η is a unitless quantity.
d. Data for **TRI**L16C from ref. 72.
e. Data for **TRI**L9C from ref. 265.
f. Data for **TRI**L16C from ref. 127.
g. Data for **TRI**L9C from ref. 125.
h. Data for **TRI**L12C from ref. 127.
i. Data for **TRI**L19C from ref. 265.
j. Data for **TRI**L12C from ref. 266.
k. Data for **TRI**L12CL16C from ref. 134.
l. Data from the crystal structure of $[Hg(II)]_S[Zn(II)(OH_2/OH^-)]_N$(**CS**L9PenL23H)$_3^{n+}$ for the trigonal HgS_3 site. Average of the three Hg-S bonds to Pen, 2.21, 2.07, and 2.41 Å[70].
m. Data from the crystal structure of $[Hg(II)]_S[Zn(II)(OH_2/OH^-)]_N$(**CS**L9PenL23H)$_3^{n+}$ for the T-shaped Hg(II) site[70].

A deeper analysis of the relationship between the peptide self-association affinity and Hg(II) (and Cd(II)) binding with a series of peptides demonstrated a linear free energy correlation with metal binding (Figure 1-9). In the case of Hg(II), this helped confirm the notion that trigonal Hg(II) complexation is a direct result of the designed peptide recognition. Further, attempts to form heterotrimeric complexes between **TRI**L2WL9C and **TRI**L2WL23C and Hg(II) did not succeed, demonstrating that control over orientation of the peptides is retained in the face of metal binding. The self-

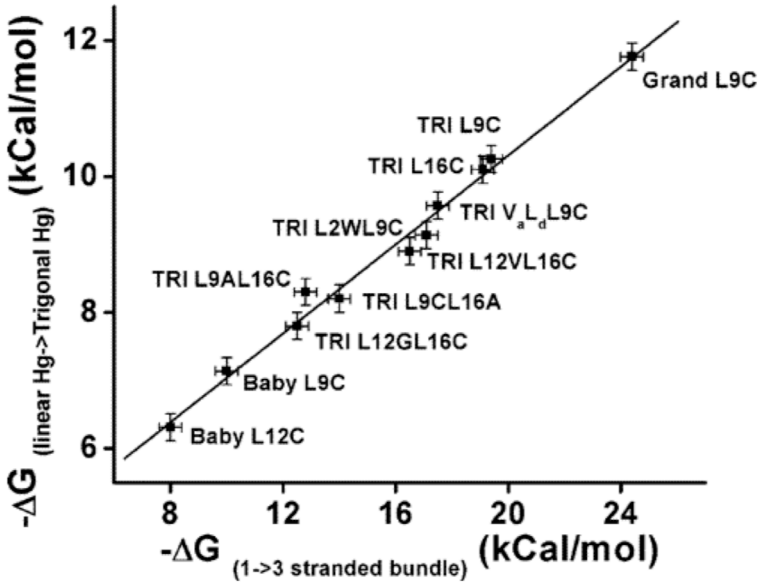

Figure 1-9. Linear-free energy correlation between folding preferences of the peptides in the absence of metal to the binding of a third strand of peptide to a divalent Hg(II)S_2 species. Figure reproduced from ref. 267.

association affinity as it relates to Hg(II) binding can be examined more deeply by considering the peptides used in the analysis. This affinity can be increased by increasing the hydrophobic contacts in the coiled coil by incorporating an extra heptad (this sequence is called **Grand** instead of **TRI**) and can likewise be decreased by either removing a heptad (to generate a sequence named **Baby**) or by substituting less hydrophobic residues for Leu around the metal site (such as Val, Ala, and Gly). At the conditions under which these peptides are typically studied (low μM concentrations), **Baby** exists as random coils in solution, **TRI** is folded into an α-helical coiled coil, and **Grand** forms an even tighter coiled coil complex. Typically, the α-helical character correlates with association of the coiled coils (amphipathic α-helices will not be well folded in aqueous solution if their hydrophobic faces are exposed), therefore, by monitoring the signal at 222 nm using CD, one can determine how well folded a given peptide is. Fitting a plot of the signal at 222 nm vs increasing concentration of guanidine hydrochloride (GuHCl) denaturant can yield the free energy of folding for the complex. As expected, when Leu residues are replaced with smaller and less hydrophobic residues,

the peptide association becomes weaker (~1-2 kcal mol^{-1} per residue per peptide). This effect is directly correlated with the affinity of Hg(II) (and Cd(II)) for each of the peptides, demonstrating that it is not the metal's geometric preference controlling the association of the coiled coils, but the free energy of folding of the peptides. This conclusion supports the theory that in biological systems, the surrounding protein matrix mediates folding and stabilization of unusual metal sites.

Trigonal planar Hg(II) complexes could be obtained in all of these designs, but the kinetics of metal binding is different and can again be related to the self-association of the peptides. Above I described how a heptad could be added to or removed from **TRI** to produce the **Grand** and **Baby** sequences, respectively. While **TRI** is well folded at a 10 μM concentration, **Baby**L9C has only 20% helicity (% helicity is the measured molar ellipticity as a fraction of the maximum ellipticity for a fully α-helical complex of the same size). However, in the presence of Hg(II), **Baby**L9C could still form a stabilized 3SCC. Initially, this challenged the notion that a preorganized site is necessary to stabilize alternate coordination geometries around metals in designed systems. To this end, a thermodynamic model was derived to explain how small, unstructured peptides could ultimately stabillize a $[Hg(II)(SR)_3]^-$ site.[53,126] First, a 2SCC is nucleated by the addition of Hg(II) to a solution of peptide because two-coordinate linear Hg(II) binding to thiolates is a highly favored process. When a third equivalent of peptide monomer is added, peptide-peptide interactions overwhelm the coordination preference of the metal to give a 3SCC final structure. Kinetic studies examining **Baby**L9C and Hg(II) binding showed that the initial binding and formation of a $[Hg(II)]_S$(**Baby**L9C)$_2$ site in a 2SCC is very fast while formation of $[Hg(II)]_S$(**Baby**L9C)$_2$(H-**Baby**L9C) with addition of more peptide is rate-limiting. This step may be described by the stepwise aggregation-deprotonation (StepAD) model in which the third helix binds to the 2SCC in either a parallel (productive complex) or antiparallel fashion (unproductive) and every time antiparallel binding occurs, the helix must first dissociate then reassociate in a parallel manner (Figure 1-10). This is followed by deprotonation of the third Cys residue at the appropriate pH to form a fully trigonal $[Hg(II)]_S$(**Baby**L9C)$_3^-$ complex. The pK_a for deprotonation depends on the location of Cys in the sequence, but is independent of the

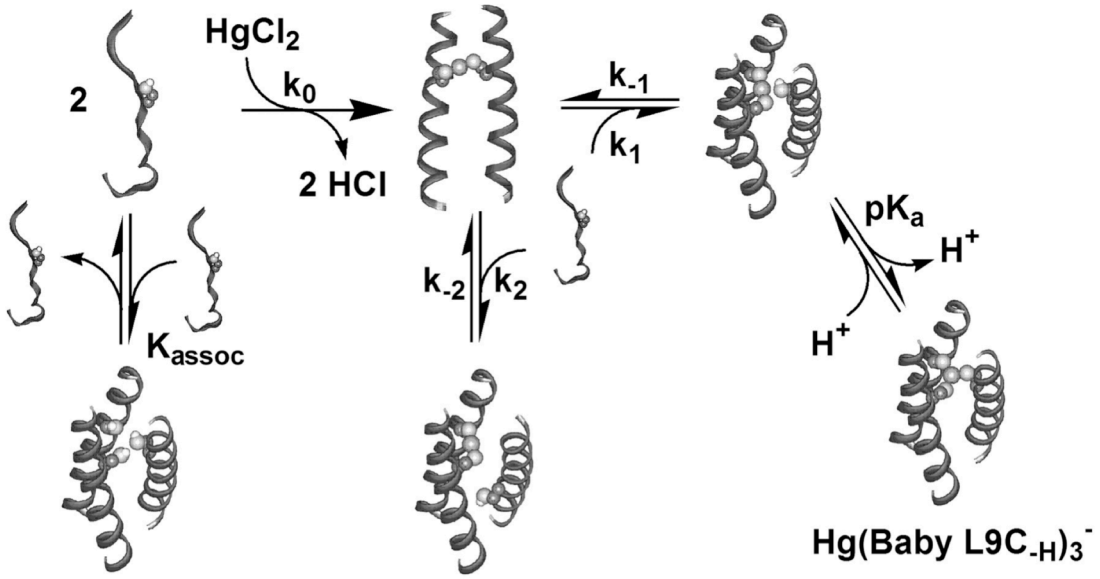

Figure 1-10. Stepwise aggregation-deprotonation (StepAD) mechanism for the encapsulation of Hg(II) by the 3SCC, (BabyL9C)$_3^{n-}$. Hg(II) reacts in a fast step to form [Hg(II)]$_S$(BabyL9C)$_2$. Formation of [Hg(II)]$_S$(BabyL9C)$_2$(H-BabyL9C) is the rate-limiting association and, depending on pH, rapidly converts to [Hg(II)]$_S$(BabyL9C)$_3^-$. Figure reproduced from ref. 126.

length of the sequence (this pK_a is 7.8 for both **a** site peptides, **TRI**L9C and **Baby**L9C while for **TRI**L12C and **Baby**L12C, both **d** site peptides, it is 8.4).

The case of Hg(II) binding to the **Baby** peptide is particularly informative as to how metals can stabilize protein folds. Since **Baby** exists at μM concentrations as unassociated and unfolded individual peptides, addition of metal is essential for proper folding. Stopped-flow kinetic analysis revealed that under all of these conditions, Hg(II) complexation to two thiolates first generates the [Hg(II)]$_S$(**Baby**)$_2$ species, which then acts as a nucleating agent for the third strand of the bundle to associate.[126] Thus, it can be concluded that the presence of the metal site stabilizes the preferred fold (3SCC) by initiating the assembly of the bundle through the non-preferred initial 2SCC. Once the system begins to fold by forming the 2SCC, it subsequently adopts the preprogramed 3SCC structure. This example illustrates that a metal site may induce a desired protein fold by stabilizing an early folding intermediate even though the final, desired fold leads to an undesired coordination environment for the metal. This work clearly demonstrates

the intricate balance and interplay between metal-binding preferences and protein folding.

The difference in the observed pK_a values for Hg(II) binding to **Baby**L9C and **Baby**L12C prompts a further examination of the difference between **a** and **d** sites. A comparison between Hg(II)-induced folding in **Baby**L9C and **Baby**L12C revealed not only different pK_a's but also a lower extinction coefficient (for the ligand-to-metal charge transfer transition representing Hg(II)-thiolate bond formation) and Hg(II)-binding affinity and a lower free energy change for Hg(II) complex formation with **Baby**L12C. A concentration-dependence study of coiled coil association revealed a higher self-association affinity for the **a** site peptide. Kinetic studies revealed a faster rate for Hg(II) binding to **Baby**L12C, likely due to this lower self-association preference (more monomers in solution can lead to more rapid nucleation). Modeling studies demonstrated the Cys conformation found in **d** site peptides is less favorable than that in **a** sites because the **d** site amino acid must twist from its normal position in the apo-bundle and orient

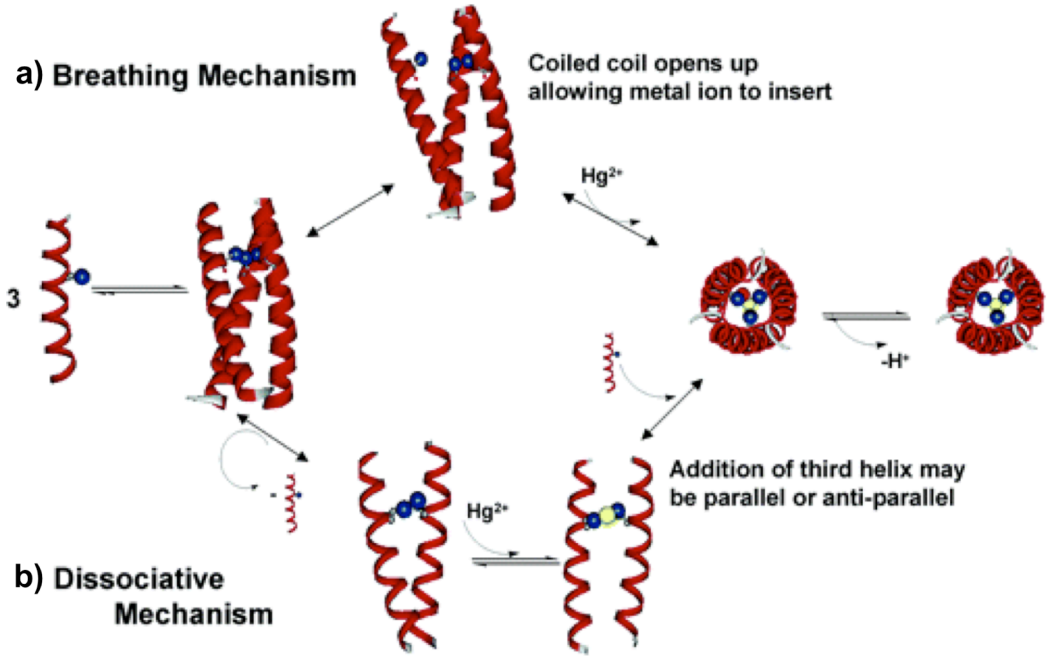

Figure 1-11. Possible mechanisms for insertion of Hg(II) into the folded peptides. Figure reproduced from ref. 53.

more towards the core, whereas the **a** site residue resembles a more preorganized metal-binding site. As discussed earlier in this section, these modeling studies have now been confirmed crystallographically with apo structures of **CS**L9C and **CS**L19C.[67]

Further kinetic studies performed on **TRI**L9C and **TRI**L12C showed the same **a** vs **d** site pattern where Hg(II) insertion into the **d** site was faster than the **a** site. However, the mechanism of metal insertion into these prefolded peptides must be different than that derived for the unfolded peptides. There are two possibilities (Figure 1-11); one is the breathing mechanism in which the 3SCC may open up and allow the metal ion to insert and the other is the dissociative mechanism, which is similar to the StepAD mechanism, in which the third helix must first dissociate and then reassociate in the proper orientation to result in a productive 3SCC. The observed biphasic behavior of Hg(II) encapsulation can be most easily rationalized with the dissociative mechanism; if some are misfolded then the step where a peptide must be lost from the 3SCC and then reassociated in the correct manner would represent a slower second phase. However, the breathing mechanism cannot be ruled out since the same pattern as was observed for **Baby** is observed for **TRI**, where the kinetics for **d** site insertion is faster. More experiments are needed to rationalize these mechanisms; however, some general conclusions for the field of protein design can be drawn from this work. One important observation is that a highly stable apo-protein may not be kinetically optimal for metal binding (another example where this is the case is described in the following Section 3a). It is likely that nature makes compromises between rapid kinetic transfer and high metal ion-protein affinity. Further, highly stable scaffolds are not necessarily required to enforce a non-preferred metal coordination environment (often required for functional metalloproteins) and may actually be detrimental to the kinetics of metal ion insertion.

3. Designing functional metalloproteins

The ultimate goal of metalloprotein design is to be able to prepare metalloproteins that can efficiently catalyze virtually any reaction as native enzymes can do for biological reactions in mild, aqueous environments with as-yet unrivalled speed and specificity. Attempts at generating catalytic proteins and metalloproteins came soon after the first

protein design papers, but until recently most have demonstrated rather poor efficiencies (if they have any reactivity at all). Many of the attempts at designing enzymes have been reviewed[7,8] and here I will focus on some of the most recent examples of functional *de novo* designed metalloproteins. In this section, I will exclude those functional examples which are Zn(II)-dependent. Given that the focus of my doctoral work is on the design of hydrolytic Zn(II) enzymes, Section 4 will describe in more detail the chemistry of Zn(II), the design of Zn(II)-binding sites and, ultimately, set the stage for the design of Zn(II)-dependent hydrolytic enzymes, the topic on which this dissertation is primarily focused.

a) De novo design of a functional diiron protein

DeGrado and coworkers have developed the Due Ferri (two-iron; DF) family of artificial metalloproteins as models for diiron and dimanganese enzymes, a diverse set of enzymes capable of catalyzing both hydrolytic and oxygen-dependent reactions.[83,84] The binuclear metal site in these proteins is bridged by a combination of oxo, hydroxo, or carboxylate donors and although the protein folds of this diverse family of metalloproteins are complex, a simple four-helix bundle houses the active sites.[15] DeGrado *et al.* have designed four-helix bundle proteins to aid in understanding how polarity, solvent accessibility, and electrostatic environment influence the properties of these proteins, with a focus on diiron sites. The work demonstrates that minimal protein models can be used to track structural changes as small as a methyl group and connect them to the enzyme's activity. Their success lies in iterative design and rigorous characterization of each model. Retrostructural analysis of diiron proteins was used to design the parent model, DF1, which is an antiparallel dimer of 48-residue helix-loop-helix (α-2) motifs that binds a di-metal cofactor near its center.[15] The helical bundle has a slight left-handed tilt as in classical four-helix bundles and antiparallel 4SCCs and its geometries may be discussed in terms of the heptad nomenclature described in Section 1c (Figure 1-12). Each helix has a single Glu residue in an **a** position which projects towards the center of the bundle. Two carboxylates bridge both metals, while the other two each interact with a single metal in a mono- or bidentate binding configuration. Two His residues are located at **d** positions, three residues away from the bridging Glu side chains

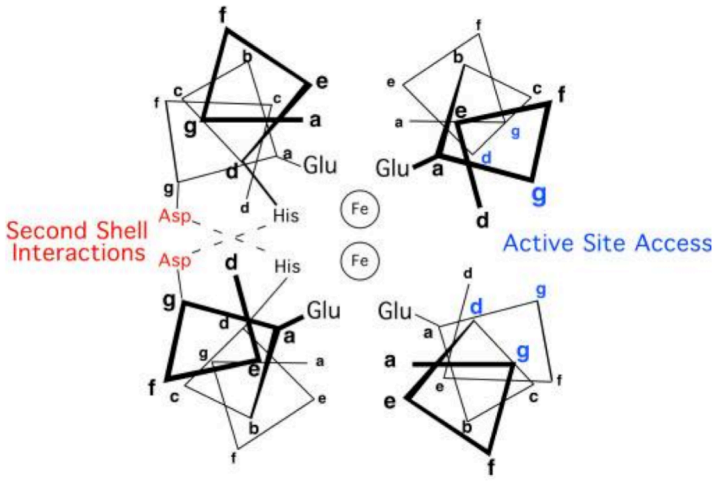

Figure 1-12. Helix-wheel representation of the antiparallel four-stranded coiled coil structure surrounding the active site of a diiron protein. Residue positions are labeled according to the heptad repeat generally applied to coiled coils. Figure reproduced from ref. 15.

and coordinate each metal in a monodentate fashion to form an EXXH binding motif. Small side chains occur at the analogous **d** positions opposite the His residues and control oxygen and substrate access. Residues in the **e** and **b** positions line the two sides of the active site (**e**/**b** interface) and the other two sides are lined by **c** and **g** residues (**c**/**g** interface). Those residues along the **b**/**e** interface play primarily a structural role and are tightly packed. Those residues at the **c**/**g** interface form hydrogen bonds to the ligand His residues on one side and on the other side define entry to the active site. The **a** and **d** positions above and below the ligand residues form the top and bottom of the active site. These are often hydrophobic and their association provides part of the driving force for association of the α-helical bundle.

DF1 was designed beginning with a *D2*-symmetric four-helix bundle of unconnected 21-residue helices and is made up of a Glu_4His_2 active site, where two Glu residues bridge the metals, the other two Glu residues are non-bridging ligands, and the Nδ of each His is bound to a metal (sequence DYLRELLKLELQLIKQYREALEYVKLPVLAKILEDEEKHIEWLETILG).[15] Additionally, second shell residues were added, which serve as geometric constraints and may be important to function. Asp35 forms a hydrogen bond with an active site His residue in a neighboring helix and Tyr7 (in a **d** position) forms a hydrogen bond with a

non-bridging Glu residue from a neighboring helix. A combination of hydrophobic and polar residues were chosen for the interfaces, helix-stabilizing polar residues were placed in the most solvent-accessible locations, and an interhelical loop was engineered in between each pair of helices. Structural analysis of di-Zn(II)-DF1[15], di-Mn(II)-DF1[135], and apo-DF1[136] revealed that the metal-binding site and surrounding secondary interactions are very similar to the intended design and largely preorganized for metal binding (Figure 1-13). Each iron is five-coordinate with a sixth vacant site lying on adjoining faces of the two metal ions, providing a potential binding site for exogenous ligands. Because the design of DF1 required the burial of six polar ionizable groups in the core of a four-helix bundle, the interior was packed with a large number of hydrophobic side chains to drive folding. These requirements for conformational stability are contrary to those for binding and enzymatic functions, as revealed by the difficulties in binding the metal cofactor (DF1 must be unfolded and refolded in the presence of metal to support binding) and the lack of any activity. In particular, analysis of the structure revealed that bulky Leu residues in **d** positions opposite the His-binding residues were likely blocking access of O_2, which would limit its capability of supporting any activity. DF1 was redesigned to carve out a larger access site by mutating each of the Leu residues to either Ala or Gly (L13A-DF1 and L13G-DF1)[135–138]. These three designs were used to examine the thermodynamic costs of carving an active site access channel, with DF1 being the most stable and the Gly derivative being the least (a total loss of ~10 kcal mol^{-1}). To confirm the presence of this designed cavity, the crystal structures of di-Mn(II)-L13A-DF1 and di-Mn(II)-L13G-DF1 were solved.[136] The X-ray structure of di-Mn(II)-L13A-DF1 revealed three dimers in the ASU, all nearly identical to those observed for the di-Zn(II)- and di-Mn(II)-DF1 structures.[138] Further, an aqueous channel is formed by the Leu to Ala substitutions and is filled by ordered water molecules and evidence of a metal-bridging dimethyl sulfoxide (DMSO) molecule. The X-ray crystal structure of di-Mn(II)-L13G-DF1 shows that with substitution by the smallest amino acid, Gly, the size of the aqueous channel increases even more.[137] There are now four independent dimers in the ASU, more water molecules present around the dimetal cofactor and no evidence of bound DMSO. More interesting is the presence of two different coordination environments around the di-Mn(II) centers (Figure 1-14). Three of the

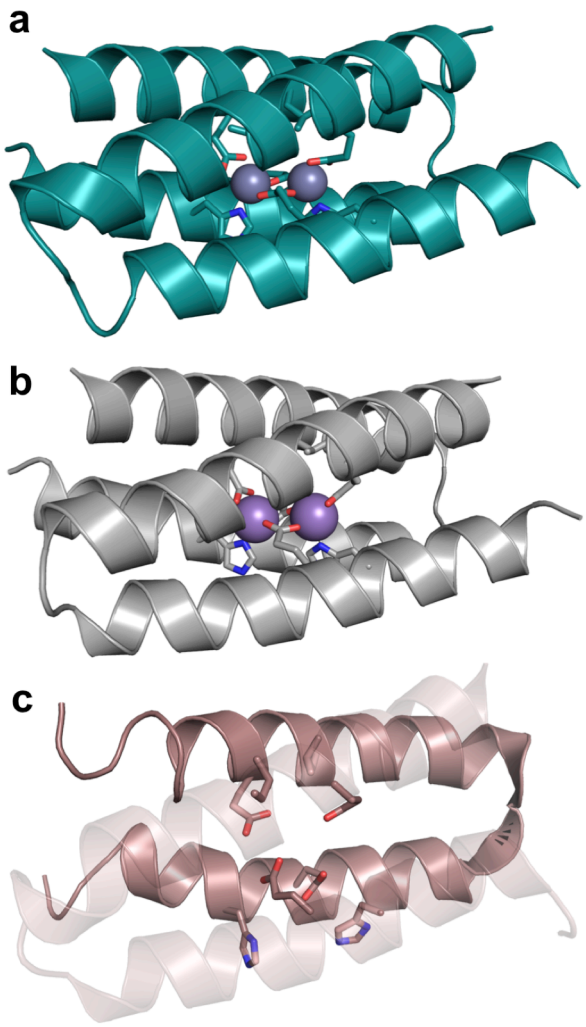

Figure 1-13. Ribbon diagrams of the structures of metal-bound and apo-DF1. (a) The X-ray crystal structure of di-Zn(II)-DF1 (pdb 1EC5)[15]. (b) The X-ray crystal structure of di-Mn(II)-DF1 (pdb 1OVR)[135]. (c) The NMR structure of apo-DF1 (pdb 1NVO)[136]. The helix-loop-helix motif in front of the ligands is shown as transparent to give an improved view.

crystallographically independent dimers in the ASU have bridging water molecules and an intermetal distance of 3.6 Å. However, the fourth has two terminal water molecules, one coordinated to each metal and *trans* to the liganding His residues. The intermetal distance in this dimer is 4.2 Å. Superimposition of the C-terminal helices (those containing the HXXE liganding site) reveal no difference in the four dimers but the N-terminal helices (those containing only a single liganding Glu residue) occupy different positions in the dimer containing the distinct metal coordination site. Particularly, two

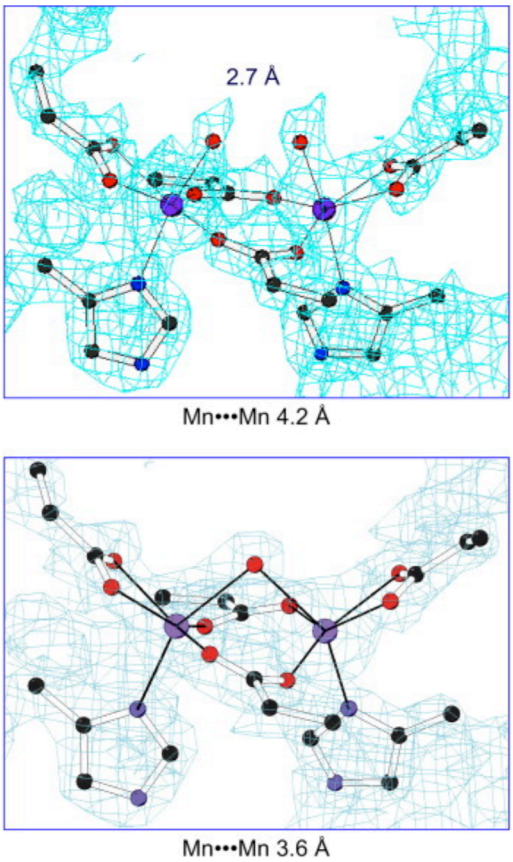

Figure 1-14. $2F_o$-F_c electron density maps of the dinuclear metal-binding site of di-Mn(II)$^{-}$L13G-DF1 in the dimer (one of four in the ASU) with terminal waters bound to each metal (top) and in one of the dimers with a bridging water (bottom). Figure reproduced from ref. 83.

copies of the N-terminal helices shift in opposite directions; this sliding helix-induced shift is what leads to the increase in the metal-metal bond length. It is postulated that this may occur in native enzymes to accommodate coordination environment changes; however, it would likely be difficult to observe in a complex, large molecule. Structural analysis of DF1 and its mutants with bound di-Zn(II) and di-Mn(II) demonstrate that the overall protein fold and general coordination environment around the dimetal site are retained with introduction of a metal site access channel. Although the DF1 constructs have proven useful for detailed structural characterization (as their stability facilitates crystal growth), they have only limited solubility and therefore cannot be fully characterized in solution. To improve the properties of DF1, some hydrophobic residues on the surface of L13A-DF1 were replaced with polar residues and a redesigned

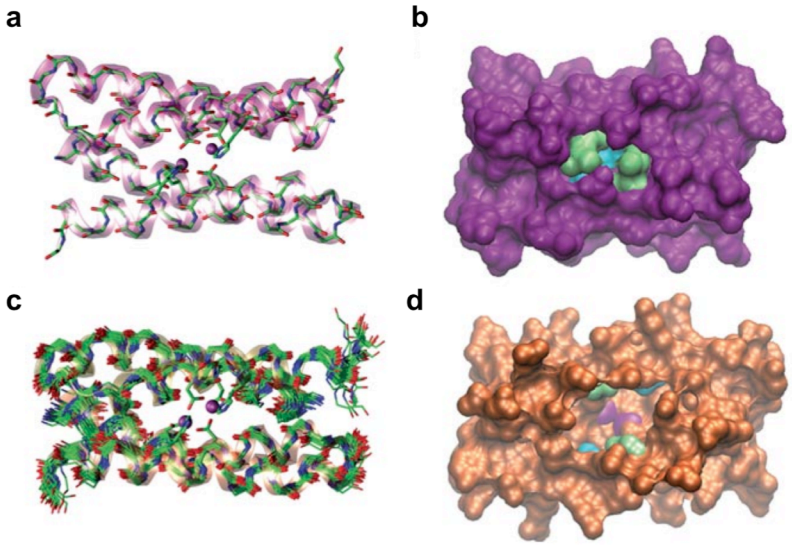

Figure 1-15. Comparison of DF1 and DF3 structures. (a) Crystal structure of di-Zn(II)-DF1 (pdb 1EC5). (b) Surface representation of the crystal structure of di-Zn(II)-DF1 to display the accessibility to the dimetal site. (c) NMR structure of di-Zn(II)-DF3 (pdb 2KIK). (d) Surface representation of the NMR structure of di-Zn(II)-DF3 to show the increased accessibility to the active site over DF1. The different residues in position 9 (lime) and 13 (cyan) are highlighted. Metal ions are colored magenta. Figure reproduced from ref. 82.

interhelical loop in DF2 and then a longer interhelical loop was incorporated in DF2t.[139–143]

Next, in searching for functionality, DFtet was designed by bridging rational design with a combinatorial approach to facilitate preparation of a variety of mutants.[144] The results obtained from the DF1 and DF2 models indicated three main aspects to be considered in order to obtain stable and functional models: interhelical turn design, residues responsible for primary and secondary ligands, and residues that define the active site access channel. In order to fully examine how systematic changes would affect these aspects, a large number of combinations of mutations would be required. DFtet was designed with four disconnected helices that could be separately prepared then combinatorially assembled. To this end, A_2B_2[144] and $AA'B_2$[145] heterotetrameric complexes were assembled and renamed DFtet. Design of these sequences implemented the use of negative protein design by destabilizing both homooligomeric folds and undesired heterotetrameric topologies. DFtet-A and DFtet-B only fold and assemble with the desired stoichiometry when mixed in a 1:1 ratio. A substrate pocket was designed into

the interior of this symmetrical DFtet that lead to the development of an O_2-dependent phenol oxidase.[6] Specifically, G4-DFtet, which had Gly residues substituted for Leu15 and Ala19 in both A chains and, as a result, the least steric restrictions of all the variants, showed rapid formation of the oxo-species with no detectable intermediates. G4-DFtet could perform the two-electron oxidation of 4-aminophenol (4AP) to benzoquinone monoamine with a 1000-fold rate enhancement over the background reaction and at least 100 turnovers. O_2 oxidizes the diferrous form of the protein, giving rise to the diferric form which reacts with 4AP. The resulting diferrous form is oxidized again to complete the catalytic cycle. The rate was decreased 2.5- and 5-fold when either Gly residue at position 19 or 15 was changed to Ala. The catalytic efficiency is therefore sensitive to changes the size of a methyl group, and demonstrates the correlation between catalytic activity and size of the active site cavity.

Table 1-3. Comparisons of the kinetics of 4-aminophenol by designed diiron proteins.

Catalyst	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (M^{-1} s^{-1})
di-Fe(III)-DF3[a]	0.045 ± 0.003	1.97 ± 0.27	23.0
G_4DF_{tet}[b]	0.022 ± 0.002	0.83 ± 0.06	25.7

a. Taken from reference 82; measured at pH 7.0 and 25° C. The protein can perform 50 turnovers over the course of one hour.
b. Taken from reference 6; pH 7.0 and 25° C. The protein performed > 100 turnovers. ~1,000-fold rate enhancement observed over the background reaction (from comparing initial rate in presence and absence of protein).

Despite this success with DFtet, its borderline stability, complex stoichiometry, and tendency to undergo ligand-exchange reactions limited full characterization. Unfortunately, DF1 is not stable enough to handle the analogous Leu15 and Ala19 mutations in DFtet (Leu9 and Leu13 to Gly) to accommodate the phenol substrate. The interhelical turn sequence was therefore modified to create more stabilizing electrostatic interactions leading to DF3.[82] The solution structure of di-Zn(II)-DF3 demonstrates that the structure remains nearly identical to the design, with the exception of a larger access channel (Figure 1-15). This site could also bind iron, undergo reversible oxidation-reduction cycles, and catalyze the oxidation of 4AP for at least 50 turnovers. The protein follows Michaelis-Menten kinetics in the presence of O_2, with an efficiency of 23 M^{-1} s^{-1}, K_M = 1.97 ± 0.27 mM and k_{cat} = 0.045 s^{-1} (Table 1-3). The designed metalloprotein could

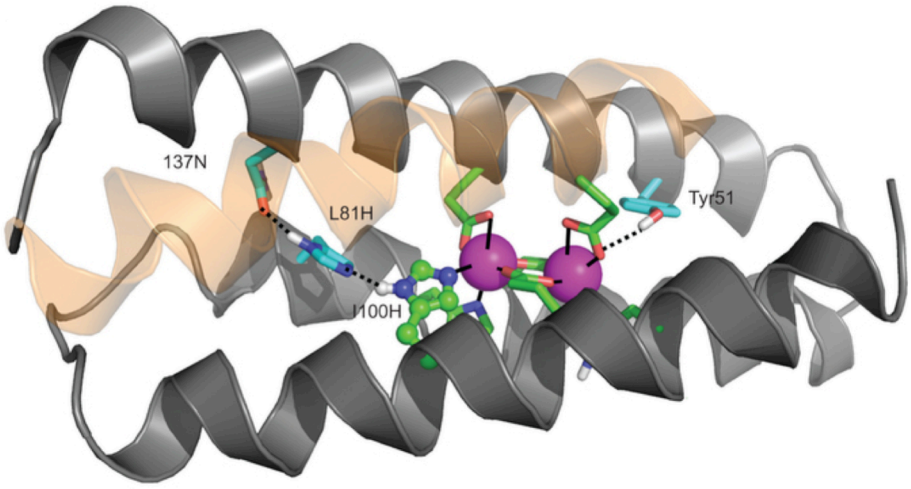

Figure 1-16. Ribbon diagram showing the structure of the 3His-G2DFsc variant (pdb 2LFD) with the added metal-binding His residue (H100) and supporting mutations (I37N, L81H). Although kinetic data is reported for the 3His-G4DFsc variant, the structure shown here contains only two Gly mutations because of increased stability during data collection (this variant also showed *N*-oxygenase activity, albeit slower than the G4-variant). Figure reproduced from ref. 81.

also catalyze the oxidation of other related substrates, such as 3,5-di-tert-butylcatechol (3,5-DTBC) (with a five-fold higher k_{cat} than that for 4AP) and *p*-phenylene diamine (expected to bind more weakly, and had a 75-fold lower k_{cat}/K_M than that for 4AP). No activity was observed for *o*-phenylenediamine. These results are all consistent with the anticipated catalytic properties of di-Fe(III)-DF3. Analysis of the crystal structure of di-Zn(II)-DF3 revealed that the L9G and L13G mutations created an active site cleft which, at its narrowest, was the width of a phenyl ring. Substitutions at the 3 and 5 positions could be accommodated in wider regions of the cavity. Thermodynamic analysis of a series of DF1 mutants versus DF3 revealed that the required functional mutations were relatively destabilizing, although the redesigned turn sequence helped to compensate for this. These requirements for activity, which lead to dramatic destabilization of the protein structure, demonstrate the tradeoff between conformational stability and catalytic function, something that is particularly apparent in metalloproteins because of the additional requirement for balance between often opposing metal-binding and functional requirements.

The success of these designs suggested it might be feasible not only to transfer an active site from a native enzyme onto a minimal *de novo* designed model, but also to alter its reactivity through a redesign approach. The target diiron enzyme for this reprogramming is AurF (a *p*-aminobenzoate *N*-oxygenase), which performs the *N*-oxygenation of anilines, and is the only structurally characterized *N*-oxygenase with a diiron center. In order to reproduce the structure of this enzyme, a third His residue needed to be incorporated into the active site. Since this would be an asymmetric substitution (single mutation), the single chain monomeric DFsc protein was used (designed by specifying 26 residues and allowing the remaining 88 to be modeled computationally).[146] The NMR structure of di-Zn(II)-DFsc confirms that the fold of the protein is retained.[147] Analysis of DFsc in the context of ferroxidase activity revealed that while it demonstrated rapid formation of a diferric site in the presence of oxygen, it was not an oxygen intermediate, as with the dimeric and tetrameric DF proteins but an off-

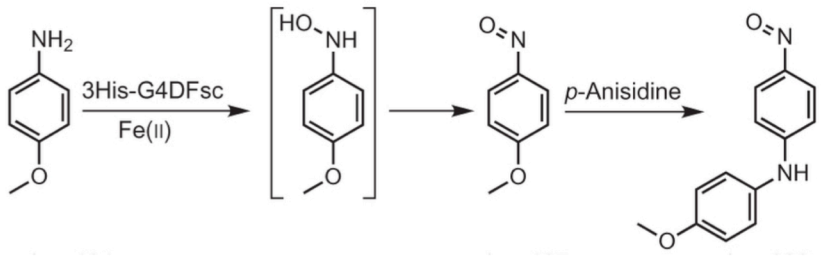

Figure 1-17. Proposed reaction scheme for *N*-hydroxylation of *p*-anisidine. Figure reproduced from ref. 81.

pathway iron-tyrosinate complex.[148,149] Expanding the substrate access channel by preparing a quadruple Ala-Gly mutant (G4-DFsc; positions analogous to those described for DF3 and DFtet) lead to a minimization of formation of this complex and rapid oxidation to form a μ-oxo-bridged diferric cluster.[81] G4DFsc also catalyzes 4AP oxidation, like other members of the DF family with open active site access channels, and with a rate on the same order of magnitude as DFsc. Subsequently, His was substituted for Ile in position 100, structurally in the same position as in AurF. To alleviate the resulting steric clashes and balance out surrounding electrostatic interactions, three more substitutions were made adjacent to the His residue to form a polar network (Figure 1-

16). Fortunately, 3His-G4DFsc is well folded in the presence of divalent metal ions but, unlike G4DFsc, does not fold in their absence, as might be expected for incorporating a number of helix-destabilizing substitutions in addition to introduction of three polar substitutions within the hydrophobic, solvent-inaccessible core. Regardless, this new construct has 4AP activity reduced to near-background levels, but can now catalyze the *N*-hydroxylation of arylamines such as *p*-anisidine (no *N*-hydroxylation reactivity is observed for G4DFsc) and other related substrates (Figure 1-17). This shift in reactivity is dramatic, especially considering that just four residues were mutated, and only one directly to the metal-binding site. Although difficult to discern in complex native diiron proteins, this study on relatively small *de novo* models suggests that it is the single His residue that directly contributes to the reactivity change. More broadly, this study is an elegant demonstration of how active site structure can be directly correlated with function, in the absence of the bulk of the surrounding protein matrix found in native metalloproteins. The loss in stability of the apo-construct is unfortunate; however, the presence of divalent metals enables characterization and reactivity studies. This decrease in stability with the introduction of function is another demonstration of the tradeoff between function and stability. The fact that metal ions stabilize the fold illustrates this interplay and how *de novo* design can be used to better understand the relationships between metal binding, protein folding, and function.

b) Dirhodium peptide catalysts for structure-selective side-chain modification

In the past few years, the Ball group has focused on the preparation of dirhodium-induced α-helices that form heterodimeric 2SCCs for a number of applications of interest to enzymatic function.[150] Particularly attractive about their strategy is that they have incorporated organometallic fragments in order to prepare bioinspired catalysts not remotely similar to what is found in nature. This group showed that they could prepare peptide complexes with dirhodium that would be stabilized or induced to form α-helices upon binding to two carboxylate groups separated by four residues and in some cases, three.[69,109] Although other peptides have been examined (and sequences as short as 9 residues could be stabilized by dirhodium binding to an *i*, *i* + 4 carboxylate motif), much

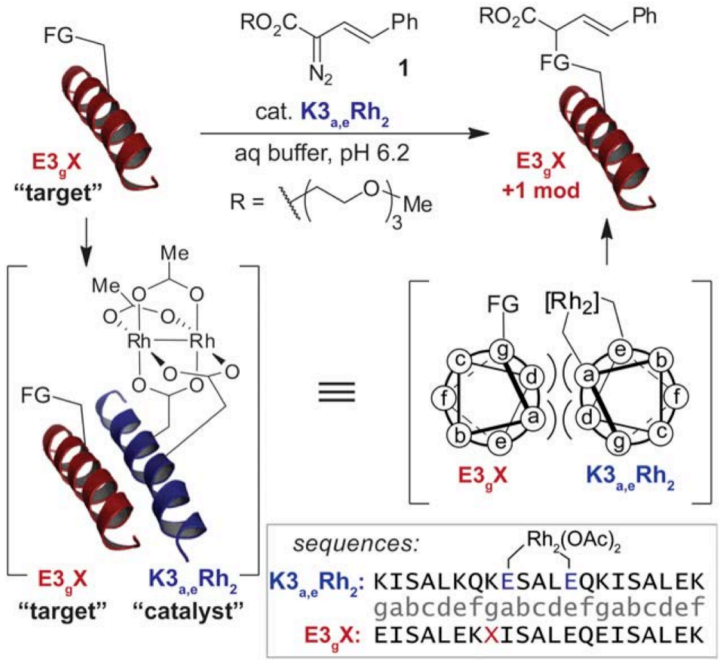

Figure 1-18. Representation of structure-selective peptide modification by a rhodium-bound peptide catalyst. 1 is the diazo reagent. X is the amino acid to be modified (Trp is discussed in the text). Figure reproduced from ref. 154.

of the work has focused on the E3/K3 heterodimeric 2SCC originally reported by Hodges.[151] The E3/K3 coil has complementary charges at the positions flanking the helical interfaces, **e** and **g**, to direct heterodimerization. By incorporating a dirhodium complex into the K3 coil at an *i*, *i* + 7 Glu motif, coiled coil formation could be inhibited, demonstrating the importance of the spacing of the Glu residues. $K3_{a,e}Rh_2$, in which Glu residues are present at *i*, *i* + 4 positions can be prepared and forms a heterodimeric coiled coil with the E3 coil. By incorporating a reactive Trp residue at the complementary **g** position of E3 ($E3_gW$) and combining this substrate with the dirhodium metallopeptide (catalyst) and a diazo reagent, a 10^3 rate enhancement for Trp modification over free $Rh_2(OAc)_4$ reagent was observed (Figure 1-18).[152,153] Further controls demonstrate that this is a site-selective molecular recognition event. Specifically, modification only occurs when the target Trp residue is placed in a complementary position of the coiled coil. When Trp is placed in a random position of the complementary strand, reactivity shows no improvement over the control reaction with free $Rh_2(OAc)_4$. Further, attempts to

modify $E3_eW$ or $K3_gW$ with $K3_{a,e}RH_2$ were not successful. This work has since been extended to include site-specific modification of a number of other amino acids (even in the presence of the more reactive Trp), some of which do not have any previous methods reported for their modification, demonstrating that molecular recognition through coiled coil formation can override the inherent chemical reactivity of the substrate.[153,154] Further, this work has been extended to demonstrate modification of larger natural targets in complex systems.[155,156] The Ball group has also shown that reactivity can be specifically inhibited through binding a His-containing substrate peptide to the metallopeptide catalyst.[157] Both of these mechanisms are important to the development of catalysts with enzyme-like characteristics including selectivity and sequence-specific inhibition. Overall, this work demonstrates the utility of designed peptide systems for performing organometallic chemistry in aqueous solutions. Related work on incorporating organometallic complexes into native proteins for this type of chemistry has also been reported[26–28], but this is the only example which exploits the known interactions in a *de novo* coiled coil system for site-specific organometallic chemistry.

4. Towards designing hydrolytic zinc metalloenzymes

The focus of my doctoral work, on designing hydrolytic zinc metalloenzymes, warrants a more thorough discussion of zinc proteins in general. Zinc is an essential cofactor in thousands of proteins. As one of the most prevalent transition metal cofactors in biological systems, it plays structural, signaling, and regulatory roles, and is found in all six classes of enzymes (most commonly hydrolases).[158] Discovery of its presence as the active metal in the hydrolytic metalloenzyme carbonic anhydrase (CA) in 1939[159] was followed by its characterization in carboxypeptidase in 1950,[160] and soon after, in enzymes of all classes.[161] In 1990, Vallee and Auld published a report analyzing the coordination spheres around Zn(II) in the available protein crystal structures, including examples from most of the enzyme classes.[162] Here, the authors introduced the spacer rule for native zinc proteins, in which at least two of the liganding residues exist within a few residues (1-3) of each other in the primary sequence, and the third was separated by a longer spacer much more varied in length (5-200 residues). This rule, in which the shorter

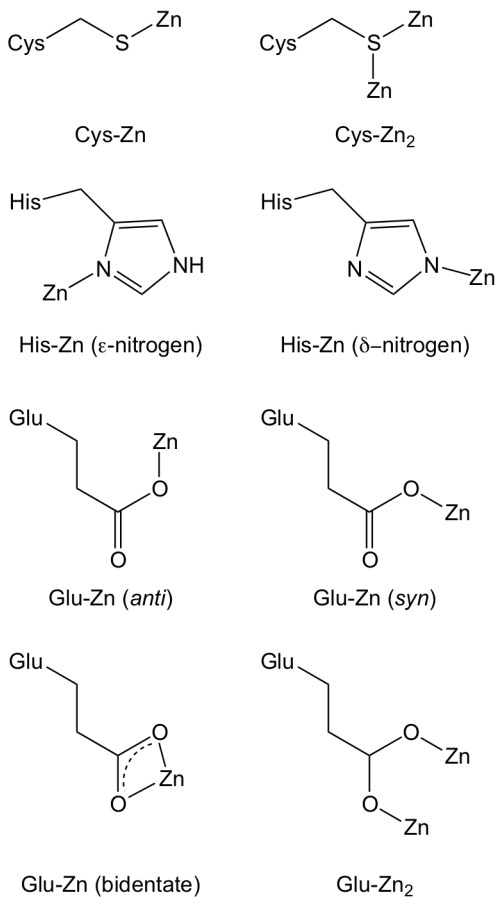

Figure 1-19. Zinc(II)-amino acid side chain binding modes as described in the text. Most zinc ligands in proteins are the side chains of cysteine (S-donor), histidine (N-donor), and glutamate or aspartate (O-donors).

spacer is proposed as a nucleus for the Zn(II) site while the longer spacer allows for some flexibility, remains rarely violated. The general guidelines for the coordination of Zn(II) in proteins have been described in a number of reports.[163] More than 99% of the time Zn(II) is coordinated by a combination of His, Glu/Asp, and Cys residues. Since Zn(II) is a borderline metal according to the hard-soft acid base theory, it can bind well to both hard (nitrogen, oxygen) and soft (sulfur) ligands. Specifically, Zn(II) can be coordinated by the N_δ or N_ε atom of the His ring, the $O_{\varepsilon 1}$ or $O_{\varepsilon 2}$ atom of Glu/Asp (*syn* or *anti*), or by the S of Cys in either an end-on or bridging fashion (Figure 1-19). More rarely, Zn(II) will also bind to the phenolate group of Tyr[164], the carboxamide oxygen of Asn or Gln[165,166], or to a protein backbone carbonyl oxygen.[167] There are several reasons for the prevalence of Zn(II) as a catalyst in enzymes. Compared to other first-row transition

metals, Zn(II) stands out because its filled *d* orbital precludes it from participating in redox reactions and allows it to function solely as a Lewis acid. This particular property makes Zn(II) an ideal metal ion for reactions requiring a redox-stable cofactor to function as a Lewis acid catalyst. As a d^{10} metal ion, Zn(II) has zero ligand field stabilization energy, so no geometry is inherently more stable than another. This lack of an energetic barrier for Zn(II) may be important for its catalytic properties, allowing for changes in coordination number throughout a catalytic cycle (from four- to five-coordinate, for example, to accommodate the intermediate) and for alterations in the reactivity of the metal ion. Additionally, Zn(II) complexes can undergo rapid ligand exchange, enhancing the ability of Zn(II) to effect a catalytic cycle through efficient product release. Although no coordination number or geometry is inherently more stable, most zinc-dependent proteins feature a resting state with a tetrahedrally coordinated zinc center (in some cases, Zn(II) is found in a five-coordinate trigonal bipyramidal geometry). This is in contrast to simple zinc complexes which usually have higher coordination numbers and are often octahedral in solution.[168] Clearly, the protein plays a role in enforcing a lower coordination number. Zinc sites in proteins have been classified into three general categories: structural, catalytic, and co-catalytic.[162,169] Structural sites have a saturated coordination sphere with all metal ligands originating from amino acids and confer stabilization on the tertiary structure. On the other hand, catalytic sites require an open coordination sphere with at least one water molecule and three or four protein ligands, and participate directly in the bond-making/bond-breaking process of a chemical reaction. Co-catalytic sites have several metals in close proximity, where one is directly catalytic and the others serve to enhance the activity. It is worthwhile to note that while these basic guidelines for Zn(II) coordination are generally true, additional principles continue to emerge for Zn(II) sites at protein interfaces, inhibitor, and transient binding sites in Zn(II)-transporting and Zn(II)-sensing sites.[104,170]

Although Zn(II) is required for the function of enzymes in all six enzyme classes, most zinc enzymes are hydrolases (aminopeptidases, carboxypeptidases, matrix metalloproteinases, phosphatases, β-lactamase, etc.). CA catalyzes the nucleophilic addition of OH^- and officially belongs to the lyase class of enzymes, but will be discussed here in the context of hydrolytic enzymes (notably, while CO_2 hydration is the

physiological reaction for CA, this enzyme can also catalyze the hydrolysis of a number of other substrates including the often used and well-studied *p*-nitrophenyl acetate, *p*NPA.[171,172]). The zinc hydrolase superfamily shares a common structural scaffold consisting of an eight-stranded β-sheet flanked by six α-helices. While the active site is located at the end of the central four β-strands, no overall conservation between the active sites exists.[173] Some members of this superfamily are monozinc enzymes (carboxypeptidase, thermolysin) while others contain co-catalytic zinc sites (aminopeptidase, β-lactamase, alkaline phosphatase). While the liganding residues around the catalytic Zn(II) ions follow the spacer rule, their identities are not conserved.[163] The active sites can be comprised of differing residues (most commonly $His_2Glu/Asp(H_2O)$ and $His_3(H_2O)$ but sometimes also $CysHis_2(H_2O)$, $Cys_2His(H_2O)$, and even $Cys_3(H_2O)$) where the residues can be located on a combination of loops, β-sheets, and α-helices. One common feature is that the active sites are generally buried within the protein structure. There are multiple evolutionarily unrelated families for CA (α-, β-, γ-, δ-, and ζ-CAs)[174,175], although all are zinc enzymes. All 16 known isoforms of the most extensively studied α-family of CAs share a conserved monomeric tertiary structure mainly comprised of β-strands.[176] Specifically, the structure is made up of 10 β-strands which form a large, twisted β-structure surrounded by six α-helices on the surface of the molecule. The active site is located at the bottom of a cavity that reaches almost to the center of the molecule where Zn(II) is bound to three His residues and a solvent molecule in a tetrahedral geometry. The His residues are located along adjacent β-strands, two with a one-residue spacer between them on a single strand (His94 and His96) and the third, separated by a longer spacer, at position 119. These examples of zinc enzymes from the hydrolase and related lyase classes clearly illustrate variability in the coordination environment, yet most hydrolases are contained within a similar overall structure centered mainly around β-sheets where, often, at least one ligand originates from a loop region or the end of a secondary structural element. Numerous differences in the protein structures encompassing hydrolytic zinc centers (some are small and monomeric whereas others are present as more complicated, multimeric structures) may tune their interactions with substrates. Determining the patterns, if they exist, between the protein folds and function of each zinc site is complicated at present, although important questions are

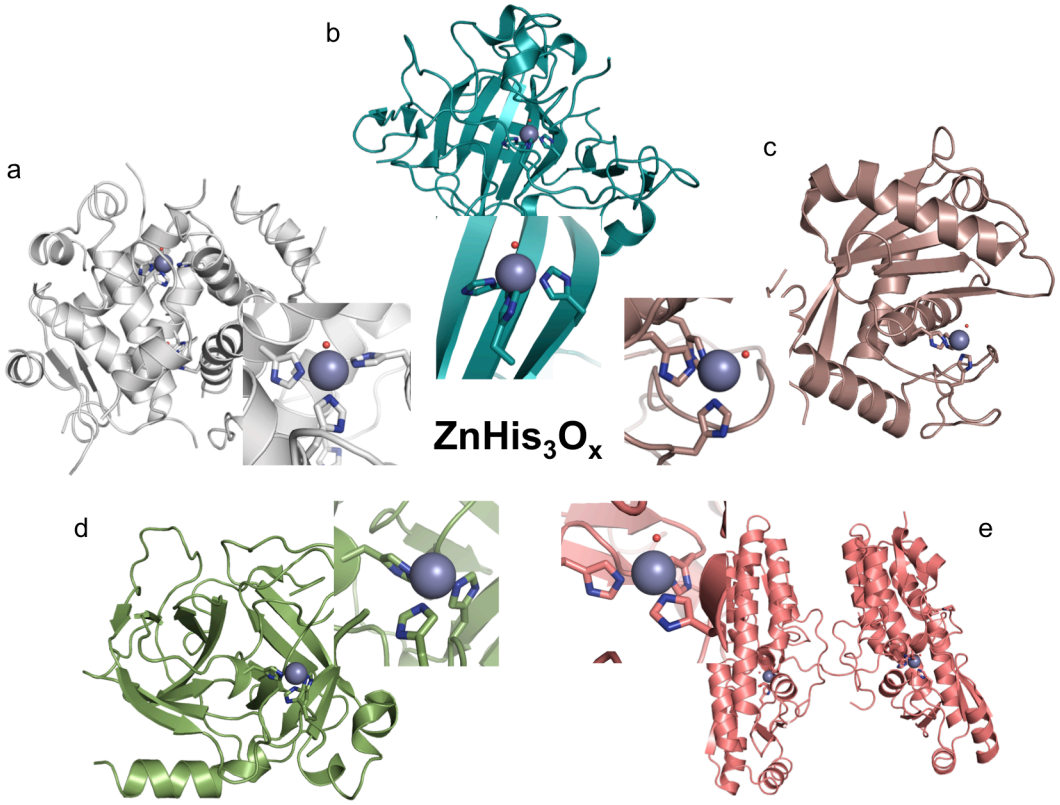

Figure 1-20. $ZnHis_3$ sites in various proteins.[104,170] a) Insulin (pdb 1AIO), in which Zn(II) organizes the hexamer with His ligands originating from three different monomers.[179] b) Carbonic anhydrase II (pdb 2CBA), in which Zn(II) forms a hydrolytic active site and each of the three His ligands are on β-sheets.[178] c) Matrix metalloproteinase adamalysin II (pdb 1AIG), in which Zn(II) forms a hydrolytic active site with two His ligands on an α-helix and the third from a loop.[268] d) Serine protease tonin (pdb 1TON), in which Zn(II) binding inhibits activity.[180] e) Zinc transporter ZnuA (pdb 1PQ4), in which the structure mediates Zn(II) mobility for transport.[181]

what the role of the protein fold is on the metal center's activity and whether it is necessary or simply a remnant of evolution.

Along similar lines, while the variability that can be achieved through just four basic amino acids and their assorted binding modes is striking, it has been demonstrated that it is not simply the first coordination sphere that differentiates one Zn(II) site and function from another. Although the importance of secondary interactions is well known[170,177], reliable guidelines such as those for the first coordination sphere around Zn(II) have yet to be well defined. For example, not only is a ZnN_3 site present in the metalloenzyme CA[178] and many matrix metalloproteinases[163], both where activation of a Zn(II)-bound water molecule to a hydroxide nucleophile occurs, it also exists as a

structural anchor in insulin[179] (each His ligand originating from a different subunit), as an inhibitor in serine protease tonin[180], and, in ZnuA[181], the ZnN_3 site transports Zn(II) in the periplasm (Figure 1-20). This variability generates a number of questions: how can a single first coordination sphere be modified by the protein environment to result in such a wide variety of metalloprotein functions? What are the ligands and structures necessary for a given function? Is the native protein fold critical for optimal function? Are there minimal changes that can be made around the metal site to alter or completely change the function? How many different functions can a single metal site perform? While direct biochemical studies on existing metalloproteins have begun to answer some of these questions, it is difficult to imagine a situation where a native enzyme can be mutated to determine the baseline requirements of the first coordination sphere in the absence of most of the surrounding protein matrix (the metal site would likely fall apart completely). Of course, this can be (and is often) approached by synthesizing small molecule models of the active sites of zinc metalloproteins, but this approach comes with a number of limitations including difficulties in working under aqueous conditions, in using the natural amino acid ligands, in stabilizing lower coordination numbers within a hydrophobic core and in assessing the influence of hydrogen bonding and electrostatics on catalytic activity.[182] Protein design is an effective yet challenging approach to recreating functional metal sites found in native enzymes. With this process, one may replicate a protein-like environment more straightforwardly than with synthetic models (using the real amino acid ligands, aqueous environment, etc.). As described earlier (Section 1a, b), the design itself can be approached in different ways (redesign or *de novo* design).[3–11,183,184]

a) Zn(II) as an attractive catalyst in proteins

Zn(II) ions serve as powerful catalysts in many hydrolase enzymes with diverse, though generally related, mechanisms.[182,185,186] The reactivity of Zn(II) in enzymes centers around a $Zn\text{-}OH_2$ moiety (Figure 1-21). In the most common basic mechanism, the Lewis acidic Zn(II) center simply promotes deprotonation of the coordinated water to generate a hydroxide species at neutral pH (as in the mechanism for CA[187–189]). For less

Lewis acidic Zn(II) centers (such as that in carboxypeptidase which has an anionic Glu residue in the coordination sphere[190]), further activation may occur by interaction with an adjacent base (such as a Glu residue) to generate the hydroxide species. A third type of basic mechanism involves displacement of the water molecule by substrate, which can then be activated (such as in 5-aminolevulinate dehydratase[191] or liver alcohol dehydrogenase[192]). There are also examples in which zinc can activate thiols towards nucleophilic attack, such as in the Ada DNA repair protein mechanism.[193,194] As will become apparent throughout the rest of this chapter, designed hydrolytic zinc proteins are generally intended to operate by the simple deprotonation mechanism and have not yet encompassed the design of nearby basic residues to assist deprotonation (a feature of many zinc hydrolases), nor have they been tested for activation of substrates such as alcohols.

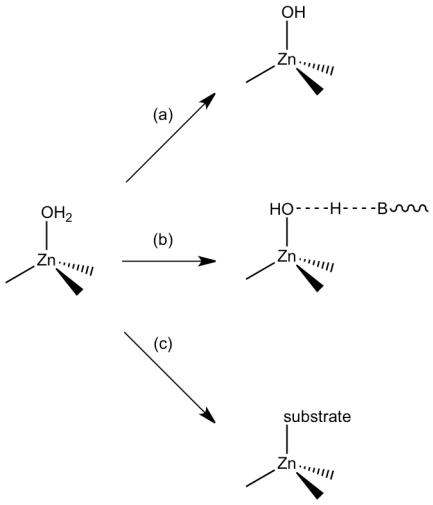

Figure 1-21. General mechanisms for Zn(II) enzymes. a) Ionization to form a Zn(II)-hydroxide nucleophile. b) Polarization with the assistance of a general base to generate a nucleophile. c) Displacement by substrate that can be subsequently activated to generate a nucleophile.

b) Designing zinc-binding proteins

De novo sites in de novo and pre-existing scaffolds

Given the importance of zinc, it is not surprising that zinc site design was among the initial metalloprotein design reports. In the earliest work, $ZnHis_3$ sites (as inspired by CA) were engineered into the *de novo* designed four-helix bundle protein, α4[76], and the redesigned antibody-like protein, minibody[90]. In the case of α4, the binding of Zn(II) to the His_3 site helped to decrease the protein's molten-globule like properties (due to a lack of conformational specificity in the core) and induce a more native-like protein fold, although no structure or binding affinity was reported (Section 1c).[49] In the minibody reported by Pessi *et al.*, a 61-residue all β-sheet structure, the dissociation constant for Zn(II) binding was estimated to have a $\sim 10^{-6}$ M lower limit (Figure 1-22a).[90] While this represents fairly strong binding, it is relatively weak when compared to native zinc enzymes which generally have dissociation constants in the nM to pM range.[195–198] A Zn(II)-binding His_3 site was also introduced into the retinol binding protein in a solvent-exposed position on the surface.[199] The dissociation constant for Zn(II) binding to this site (36 ± 10 nM) is somewhat stronger and closer to the range for native zinc proteins. The geometry, based on competition with chelating ligands, is likely octahedral, with three open coordination sites.[200] Soon after, a scorpion toxin, charybdotoxin (37 amino acids long and comprised of a short α-helix on one face and an antiparallel triple-stranded β-sheet on the opposite with three stabilizing disulfide bonds in the interior), was used as

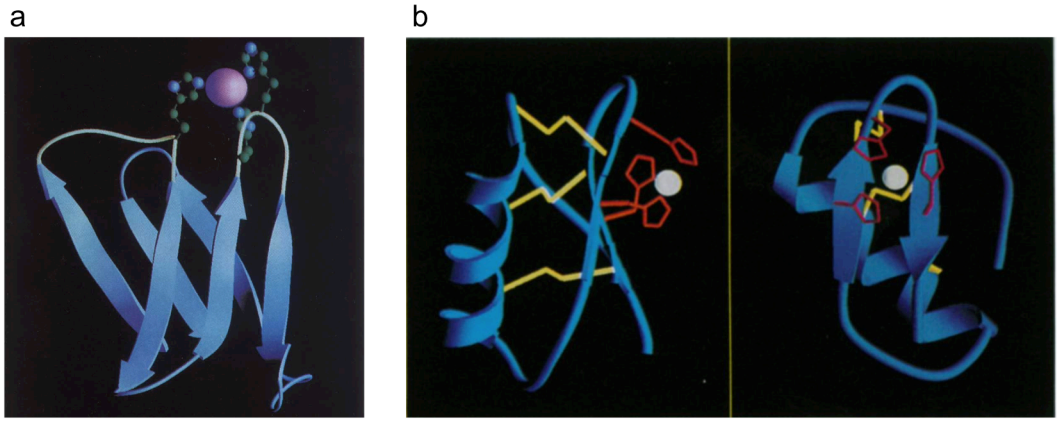

Figure 1-22. Designed $ZnHis_3$ sites. a) Modeled structure for the design of the minibody and its predicted metal-binding site (His_3). Reproduced from ref. 90. b) Model of the redesigned scorpion toxin charybdotoxin (orthogonal views) with a His_3 metal-binding site. Disulfide bonds are shown in yellow. Metal-binding ligands are shown in red. Reproduced from ref. 201.

a small scaffold for the incorporation of a His_3 site (Figure 1-22b).[201] Nine mutations were made to incorporate the metal-binding site, including engineering in three His residues and including the Gln and Glu residues found in CA which form hydrogen bonds to two of the primary His ligands. While no X-ray crystal or NMR structures are reported, CD and ^1H NMR indicates the structure is largely retained relative to the parent sequence. The dissociation constant for Zn(II) binding is 5.3 (±0.4) x 10^{-6} M and, although no catalytic activity was reported, the design demonstrates the achievement of a stable, yet solvent-exposed metal-binding site at the surface of a miniprotein.[197,202] Table 1-4 summarizes the known dissociation constants for many of the designed Zn(II)-binding sites that are reported.

The *de novo* designed four-helix bundle, α4, was also used as a scaffold for the incorporation of a closed-sphere $ZnCys_2His_2$ site (Zα4), much like the structural site found in zinc fingers (ZFs) (Figure 1-23a).[80,203] Co(II) was used as a spectroscopic probe to demonstrate successful design of a tetrahedral site. Mutants of this design, in which one ligand was removed at a time (by substitution with Ala) in order to generate an open coordination site, were also tetrahedral, although the Co(II)-binding affinities were all decreased by at least one order of magnitude. Overall, the Cys residues were concluded to be stronger determinants for Co(II) binding than His, which is not unexpected since the negative charge on the Cys thiolate residue is known to provide a stronger binding contribution.[204–207] Klemba *et al.* reported the redesign of the B1 domain of IgG-binding protein G (a 56-residue protein with a four-stranded β-sheet crossed by a single α-helix) to incorporate a closed-sphere His_3Cys site for binding Zn(II) (Figure 1-23b).[208] Co(II) substitution confirmed the presence of a tetrahedral site and, based on competition, Zn(II) binding was estimated to be on the order of 10^{-9} M. There are also numerous other studies involving the design of a variety of peptide ligands for the preparation of metal sites meant to mimic ZF proteins (coordination environment $Zn(Cys)_{4-x}(His)_x$ where x = 0, 1, or 2), which are beyond the scope of this chapter.[209,210] Overall, these examples of designed Zn(II) sites with four protein ligands generally display higher affinities than the designed sites with three protein ligands discussed above. It is worth noting that no strict correlation exists with the number of ligands bound in the first coordination sphere for native zinc enzymes.[104,195,198] This is likely due to the incorporation of secondary

Table 1-4. A comparison of Zn(II)-binding affinities to selected designed proteins.

Design	Coordination Site	K_d (Zn(II)) (M)	pH	ref
Minibody	His_3	$\sim 10^{-6}$	na	a
Retinol binding protein	His_3	36 (±10) x 10^{-9}	na	b
Charybdotoxin	His_3	5.4 (±0.4) x 10^{-6}	6.5	c
IZ-3aH	His_3	9 x 10^{-6}	7.0	d
IZ-3adH	His_6	23 (±2) x 10^{-6}	7.0	e
Zα4	His_2Cys_2	2.5 x 10^{-8}	7.5	f
Zβ1M,A,L	His_3Cys	$\sim 10^{-9}$	7.5	g
Trx[ZS].C, F	His_2Cys_2	~ 10 x 10^{-9}	7.5	h
Trx[ZS].A, B, E	His_2Cys_2	~ 1 x 10^{-6}	7.5	h
BABZ5	His_2Cys_2	2.2 x 10^{-6}	7.3	i
IGA	Cys_4	0.4 x 10^{-15}	pH 8.0	j
		4 x 10^{-12}	pH 7.0	j
GGG	Cys_4	18 x 10^{-18}	pH-independent	k
		0.5 x 10^{-12}	pH 7.4	k
	$HisCys_3$	0.7 x 10^{-15}	pH-independent	k
		0.5 x 10^{-12}	pH 7.4	k
	His_2Cys_2	40 x 10^{-15}	pH-independent	k
		0.5 x 10^{-12}	pH 7.4	k
αββ Zn(II)-hydrolase	His_3	800 x 10^{-6}	na	l

a. Taken from reference 90.
b. Taken from reference 199.
c. Taken from reference 201.
d. Taken from reference 117.
e. Taken from reference 110.
f. Taken from reference 80.
g. Taken from reference 208.
h. Taken from reference 212.
i. Taken from reference 91.
j. Taken from reference 214.
k. Taken from reference 207.
l. Taken from reference 250.

structure interactions and demonstrates the utility of protein design to uncover features of metal binding that would otherwise be difficult to determine in a native system.

Metal-mediated stabilization of protein structures

There are a number of examples where designed Zn(II)-binding sites have resulted in enhanced folding of partially or completely unfolded peptide structures.[88,91,110,117,207,211–214] As described in Section 2a, a His_3 site was introduced into a

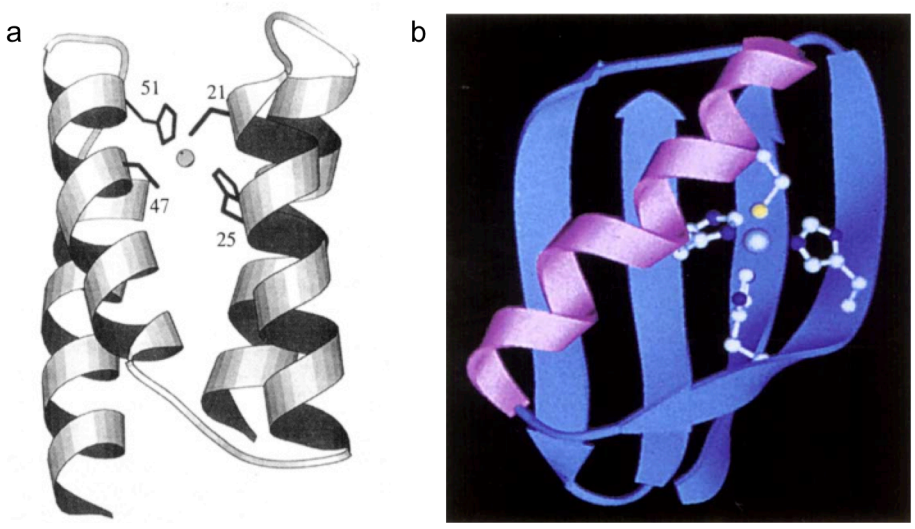

Figure 1-23. Designed $ZnHis_xCys_x$ sites. a) Model of Zα4. The side chains (clockwise from top right) are Cys21, His25, Cys47, and His51. Reproduced from ref. 203. b) Model of the metal-binding site in the B1 domain of *Streptococcal* protein G. The ribbon diagram represents the designed Zβ1M. The side chains around the metal-binding site are His16, His18, His30, and Cys33. Reproduced from ref. 208.

de novo designed 3SCC, IZ ($YGG(IEKKIEA)_4$) to generate IZ-3aH which folds in the presence of Zn(II) (K_d = 9 μM), although no X-ray crystal structure was obtained.[117] The same group reported a similar 3SCC, but with six His residues available for binding in an octahedral geometry with a K_d for Zn(II) binding of 23 ± 2 μM.[110] These affinities remain in the same range as those described above with three protein ligands (Table 1-4).

Wisz *et al.* reported the semi-automated design of a series of $ZnCys_2His_2$ centers into the non-metalloprotein thioredoxin (Trx, using the program Dezymer[215]) (Figure 1-24a).[212] Dezymer was used to identify sites where appropriate amino acids could be engineered to create a metal-binding site with predetermined geometry, in this case tetrahedral. Visual inspection identified seven variants with buried potential metal sites and few steric clashes with the backbone. Five of seven experimentally characterized mutants bound Co(II) in a tetrahedral geometry with strong dissociation constants for some (2-4 μM) and weak for others (150-250 μM). Zn(II) binding (measured by displacement) was 2-3 orders of magnitude stronger than Co(II) (nM to μM). The strongest K_d's were for structures in which the apo form was stable and metal binding

further increased the stability (Trx[ZS].C, F). In two cases (Trx[ZS].A, B), the apo form was unstable and metal binding induced folding (Co(II) K_d's ~150 μM). In the fifth case (Trx[ZS].E), the apo form was stable but metal binding decreased the stability. The weak metal-binding affinity for this variant results from clashing steric interactions within the metal coordination sphere. Thermodynamic analysis revealed that all variants bound metal to the unfolded state and the observed metal-mediated stability is a consequence of differential binding to the folded native state and unfolded structures, although the dominant factor for binding to the unfolded state is unclear. This work demonstrates how the design of a family of similar sites within the same scaffold can lead to an understanding of how both the intrinsic properties of the metal center and the surrounding protein matrix can affect the metal-binding affinity (and stabilization). In this case, the metal dictates geometry and the thermodynamic price paid by the protein to accommodate the metal is reflected in the binding constant. Important considerations for designing metal-binding sites are that packing interactions in the binding site should be carefully considered to avoid costly reorganization energies and binding to the unfolded state should be destabilized (negative design).

Another example which examined the design of a $ZnHis_2Cys_2$ site in several locations of a protein (*de novo* designed DS119 with βαβ structure) was reported by Zhu

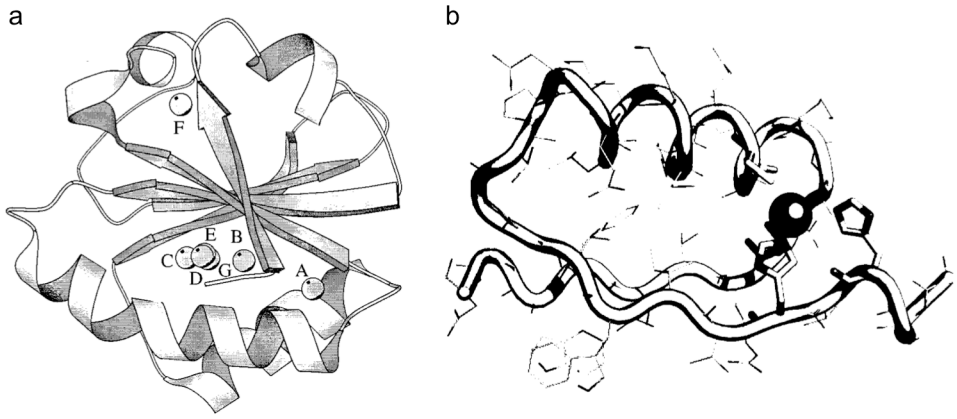

Figure 1-24. Designed $ZnHis_2Cys_2$ sites. a) The distribution of the designed Cys_2His_2 sites in thioredoxin (letters identify each design). Reproduced from ref. 212. b) Structure model of BABZ5 protein. Zn(II) is indicated as a sphere and coordinating residues (two His and two Cys) are shown as sticks. Reproduced from ref. 91.

et al (Figure 1-24b).[91] It was found that the Zn(II) sites designed in the core generally destroyed the folding even in the presence of zinc, while those at flexible terminals or loops displayed Zn(II)-induced aggregation, with K_d's in the 2-20 μM range, however, the binding stoichiometry was 1:1 in only one case (BABZ5, K_d = 2.2 μM). These observations result from a balance between the energy provided by metal binding and the cost of folding the protein structure. In the cases where the engineered ligands destroy the core-packing of the protein, energetic contributions from the binding of Zn(II) are not enough to recover the fold.

Overall, these design studies where metal binding induces or further stabilizes a protein fold demonstrate a connection between the binding constant and the energy required to fold a given protein. The formation of metal-ligand bonds provides a favorable enthalpic contribution to the free energy of protein folding. There are also entropic contributions from the release of water molecules from the binding site and the solvated metal upon binding. In an effort to separate protein-protein and metal-protein contributions towards the free energy of folding, Reddi *et al.* reported the thermodynamic analysis of Zn(II) binding to a Cys_4 site in a minimal, unstructured, 16-mer peptide, **GGG** (a variant of **IGA** with the sequence NH_2-KLCEGG·C*GG*CG*G*C·GGW-$CONH_2$).[213] Petros *et al.* had initially reported the design of a structural ZnS_4 binding site in a *de novo* designed 16-amino acid peptide ligand, **IGA** (sequence H_2N-KL*C*EGG-*C*IG*C*GA*C*-GGW-$CONH_2$)[214] and found apparent dissociation constants for Zn(II) at pH 8.0 of 0.4 fM and at pH 7.0 of 4 pM (similar to that for a redesigned ZF, CP1-CCCC, 1.1 pM[216]). These are well within the range observed for natural zinc proteins like metallothionein (0.1 pM at pH 7.0)[198], yet (at pH 7.0) weaker than those of the zinc sensor proteins ZntR and Zur (1.5 fM and 1.1 fM, respectively).[195,217] It was suggested from this work that the formation of the $ZnCys_4$ site provides -22.1 kcal mol^{-1} driving force for protein folding. The goal of the work using **GGG** (designed to more closely resemble structural Zn(II) sites than **IGA**, which was originally based on a [4Fe-4S] binding motif) was to decouple the metal binding and protein folding events by choosing a structure in which the free energies of the folded and unfolded states are identical. The pH-independent dissociation constant for Zn(II) binding to **GGG** is 18 aM, indicating that a $ZnCys_4$ site can provide up to -22.8 kcal mol^{-1} driving force for protein structure

stabilization (Figure 1-25).[207] Further, the majority of the driving force is due to dehydration of the metal and binding site so the process is largely entropy-driven. Zn(II)-$(Cys)_3His$ and Zn(II)-$(Cys)_2(His)_2$ sites were also examined with similar results and dissociation constants of 0.7 fM (and -20.7 kcal mol^{-1} contribution to protein stability) and 40 fM (-18.3 kcal mol^{-1} contribution), respectively. The observation that the Cys thiolate is better ligand for Zn(II) than His is again consistent with literature precedent.[204–207] However, it should be noted that at physiological pH (7.4) the contributions from each of the differing coordination motifs are equal, due to proton competition with the thiolate ligands (K_d = 0.5 pM) and there is a -16.8 kcal mol^{-1} energetic contribution to the protein-folding driving force. In natural proteins, which can have weaker binding affinities for the same coordination motifs, the difference is due to the loss of Zn(II)-binding energy to the protein folding process.

The work described in this section, although focused on structural zinc sites, is important in the context of my dissertation because it highlights several aspects of metal binding and protein folding that must be considered when designing zinc metalloenzymes. The energetic cost of protein folding should be both minimized in the apo form and in the metal-bound form to minimize the reorganization requirements upon metal binding. Next, the binding of metal should be optimized towards the folded state rather than the unfolded state while keeping the structure flexible enough to allow for metal binding. Although very high binding affinities may be achieved for unfolded structures such as designed peptides **IGA** and **GGG** and natural ZFs, catalytic zinc sites

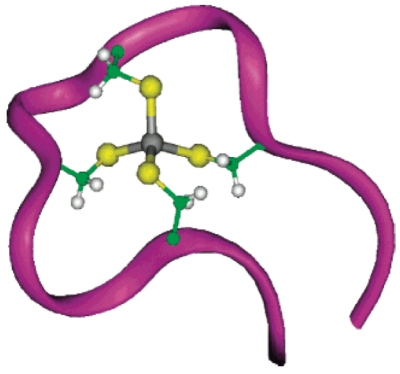

Figure 1-25. Molecular model of Zn(II)-GGG complex with a Cys_4 site rendered using Biosystem Insight II. Figure reproduced from ref. 213.

will likely require further structural contributions from a larger protein matrix for the stabilization of unsaturated coordination spheres and substrate binding. Additionally, while Cys residues can clearly provide more favorable Zn(II)-ligand interactions, their effects on the Lewis acidity of the Zn(II) center must be considered and will be discussed below. Therefore, designing catalytic zinc sites will require striking a balance between strong enough Zn(II) binding to promote catalysis (the affinities achieved using **IGA** and **GGG** variants are certainly higher than those observed for effective native Zn(II) enzymes) and formation of well-folded protein structures.

Metal-mediated stabilization towards the preparation of designed multimeric structures

Another approach to the design of structural metal sites (and, in some cases, catalytic sites) in proteins that is gaining ground in recent years is to use the binding energy provided by metals to direct the folding and assembly of multiple protein subunits and even extended nanostructures. Although metalloprotein design is typically associated with incorporating stable metal sites into protein cores, metal-protein surface interactions are at least equally significant and metal sites are often located at the interfaces of multimeric protein complexes. Additionally, the formation of metal sites at protein interfaces has been proposed as a possible evolutionary route towards efficient metalloenzymes. Therefore, gaining a full understanding of how this might occur through a design approach should prove very important. This topic has been reviewed[36,37,218], but the strategy has since expanded to include the generation of periodic protein arrays[38] and, as will be mentioned below and discussed in Chapter 2, in the design of a hydrolytic zinc metalloenzyme[18,39]. Briefly, the Tezcan group has focused on using metal-coordinating motifs on the surface of a monomeric protein, cytochrome cb_{562}, to engineer novel protein-protein interaction interfaces in order to control the assembly of proteins. The initial model system, MBPC-1 (metal-binding protein complex), has two bis-His motifs on its surface which, upon addition of equimolar amounts of divalent metals (Zn(II), Cu(II), Ni(II)), forms discrete multimeric structures according to the stereochemical preferences of the metal ions.[36,219,220] Subsequent computational redesign using RosettaDesign[221] optimized the hydrophobic packing between interfaces and included some additional hydrogen-bonding and salt bridge interactions to convert the Zn(II)-

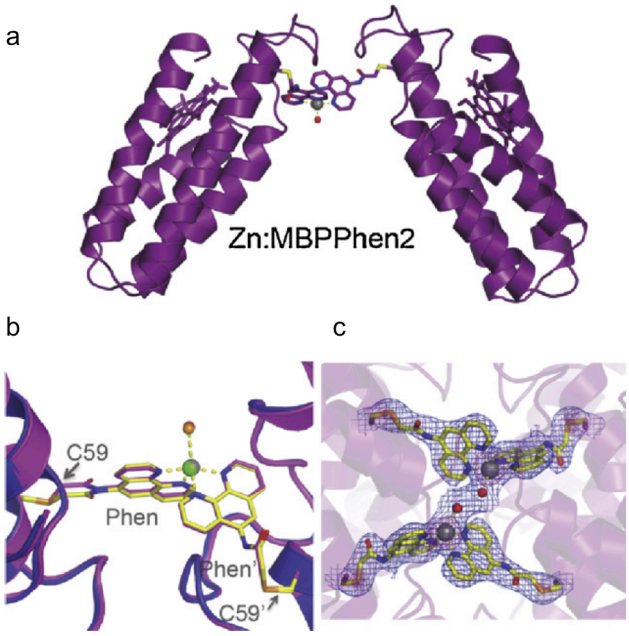

Figure 1-26. a) Crystal structure of the Zn:MBPPhen2 dimer (3MNK).[223] b) Superposition of Ni:MBPPhen2 (yellow) and Zn:MBPPhen2 (magenta) metal centers. c) Close-up view showing the proximity between the coordinatively unsaturated metal centers in the asymmetric unit of the Zn:MBPPhen2$_2$ structure. $2F_o$-F_c electronic density map is contoured at 1.2 σ. The dimer is formed both in solution and the solid-state. Reproduced from ref. 37.

mediated tetramer into a structure that could self-assemble even in the absence of metals.[222] Tezcan also successfully incorporated non-natural chelates onto the surfaces to generate not only coordinatively unsaturated metal sites, but also a site that enforces a non-preferred geometry on a metal due to the steric bulk of the chelates and lattice packing arrangements that hold the metal centers in close proximity (Figure 1-26).[223] This is of particular importance to the design community since an efficient metalloenzyme often requires coordinatively unsaturated metal sites with strained metal geometries. Achieving this requires a structure that strikes the delicate balance between thermodynamic stability at the metal site and over the whole protein structure.

Reengineering pre-existing metal sites for hydrolytic activity

Beyond the introduction of Zn(II)-binding sites into both *de novo* and existing protein scaffolds through the variety of approaches discussed above, existing Zn(II)-binding sites can be redesigned to introduce new functions or alter existing functions. The

approaches have ranged from single ligand substitutions (converting ZF structural sites to hydrolytic sites[224,225]) to more extensive redesign approaches (for example, converting glyoxalase II to a β-lactamase[33]). Additionally (although I will not go into extensive detail here), there are several examples of using Zn(II)-binding sites to control or alter different protein functions. These studies range from the design of Zn(II)-binding sites for inhibition[226,227] or control of specificity of protease activity in trypsin[228–230] to Zn(II)-induced conformational changes for sensing[231–233] and control of various protein functions[234,235].

While there are many examples of redesigned ZF peptides, here I will focus on those in which the goal is to prepare a hydrolytic site.[224,225] One route to the design of catalytic Zn(II) sites is to take an existing structural site and remove a ligand to create an open coordination sphere for binding and activation of external ligands. This was first attempted by Merkle and coworkers, who reported a truncated Zn finger peptide, CP1-C4 (CP1: PYK*C*PE*C*GKSFSQKSDLVK*H*QNT*H*TG) in which the last four residues (including a His) were removed, to create a peptide which binds Co(II) tetrahedrally.[236] While evidence of an open coordination site was obtained by examining the spectra of Co(II)-peptide complexes with external ligands, neither Zn(II) or Co(II) complexes displayed any hydrolytic activity, a result proposed to be due to the high pK_a expected for a thiolate-rich site with no secondary interactions. Later, Nomura *et al.* reported the first successful attempt at redesigning the structural zinc center in a ZF (wild-type sequence

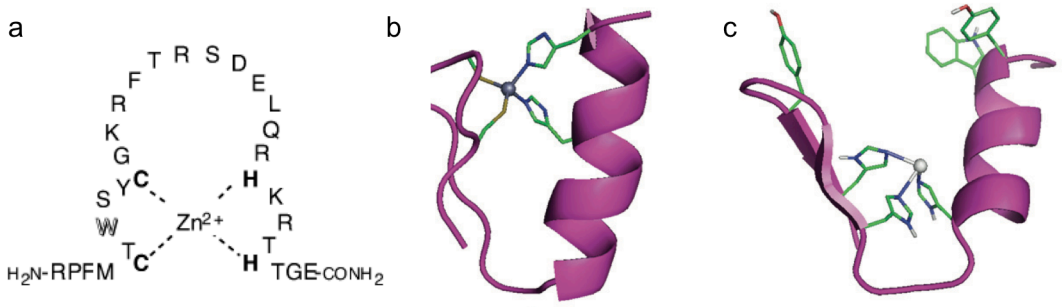

Figure 1-27. ZF peptides. a) Amino acid sequence of ZF parent peptide. Reproduced from ref. 224. b) Natural ZF fold (αββ structure), and c) ZF αββ fold modified as a metallo-hydrolase. Both b) and c) reproduced from ref. 250.

RPFM*C*TWSY*C*GKRFTRSDELQR*H*KRT*H*TGE, zf(CCHH)) into a hydrolytic Zn(II) site by preparing a series of ZF mutants in which one of the coordinating residues was substituted with a non-coordinating residue (Ala or Gly) (Figure 1-27a).[224,237,238] For those mutants retaining at least one Cys in the coordination sphere, Co(II)-substitution indicated four- or five-coordinate geometries. The highest activity (*p*NPA hydrolysis) was observed for zf(CCHA), which has a sequence similar to the truncated ZF CP1-C4 and four-coordinate geometry (with an open site) and a second order rate constant of $0.568 \pm 0.0228\ M^{-1}\ s^{-1}$. The Zn(II) center is located in a more hydrophobic environment in zf(CCHA) than in CP1-C4[236] due to the presence of the C-terminal residues, which can affect both substrate binding and also the acidity of the Zn(II)-bound water molecule.[239] All of the other mutants with at least one Cys in the coordination sphere were five-coordinate, including two vacant sites. Generally, ZFs of the CHH-type displayed higher activities than those of the CCH type, probably because the Cys residue can decrease the Lewis acidity of the Zn(II) center via its electron-donating ability and, consequently, reduce the activity (Table 1-5). To this end, the authors prepared several HHH-type ZFs, which should bind as neutral ligands (as in CA) and help maximize the Lewis acidity. As expected, increased activity was observed relative to the Cys-containing ZFs (up to $0.966 \pm 0.0492\ M^{-1}\ s^{-1}$ for zf(HHHH)). While only rarely seen in nature (for example, a structural site in the HAP1 transcriptional factor[240]), a $ZnHis_4$ site has also been observed in a related redesigned ZF protein[241] and in the MBPC-2 system reported by Salgado *et al*.[242], but no activity was reported for either. Co(II)-substitution of all of these, however, indicated six-coordinate geometries, with at least two vacant sites, although, given the preference of Co(II) for octahedral geometries, lower coordination numbers for Zn(II) cannot be ruled out. The hydrolysis of *p*NPA as performed by these ZF mutant peptides is pH-dependent with pK_a values ranging from 6.3 to 7.6 that increase as expected based on the Lewis acidity of the Zn(II) center. The activity of the His_4 site falls within the range of those reported for synthetic small molecule model complexes in mostly aqueous conditions[243–247], yet remains almost 3000-fold slower than CAII.[171,248] zf(HHHH) also displayed nuclease activity towards substrates bis(*p*-nitrophenyl) phosphate and supercoiled plasmid DNA (pUC19GC).[249] Overall, this work demonstrates the

Table 1-5. Coordination environments and *p*NPA hydrolysis rate constants of zinc finger mutant peptides.[a]

ZF variant	Coordination # (# vacant sites)[b]	*p*NPA hydrolysis rate constant (M^{-1} s^{-1})
Zn(II)-zf(CCHH)	4 (0)	0
Zn(II)-zf(CCGH)	5 (2)	0.218 ± 0.0085
Zn(II)-zf(CCAH)	5 (2)	0.232 ± 0.0051
Zn(II)-zf(CCHG)	5 (2)	0.351 ± 0.0182
Zn(II)-zf(CCHA)	4 (1)	0.568 ± 0.0228
Zn(II)-zf(GCHH)	5 (2)	0.399 ± 0.0014
Zn(II)-zf(CGHH)	5 (2)	0.458 ± 0.0021
Zn(II)-zf(AHHH)	6 (3)	0.478 ± 0.0057
Zn(II)-zf(HAHH)	6 (3)	0.497 ± 0.00058
Zn(II)-zf(HHAH)	6 (3)	0.370 ± 0.0289
Zn(II)-zf(HHHA)	6 (3)	0.443 ± 0.0147
Zn(II)-zf(HHHH)	6 (2)	0.966 ± 0.0492

a. Taken from ref. 224. In 20 mM HEPES buffer (pH 7.5), 0.1 M NaCl, 3.5% acetonitrile, 25 °C.
b. Based on Co(II)-substituted UV-Vis absorption studies.

importance of having primarily neutral ligands in the coordination spheres of Zn(II) centers in hydrolytic enzymes.

In another study, a ZF protein (αββ fold) is targeted for computational redesign as a hydrolyase (Figure 1-27c).[250] In this design method, CYANA and then IDeAS were used to design a linker between D-amino acid-nucleated secondary structures (α-helix and two β-hairpins) and to optimize the sequence, respectively. The resulting small 21-mer sequence with αββ structure bound Zn(II) with K_d ~800 μM and ΔG^o of ~ -17.2 kJ mol^{-1}. Therefore, Zn(II) binding to this His_3 site is weaker than for any of the other designs discussed above (Table 1-4). Hydrolysis of *p*NPA was observed with an initial rate of 103.45 (±0.25) x 10^{-9} M s^{-1}, a 45-fold enhancement over the background at pH 7.0 and 25 °C. However, enzyme-like hydrolysis with substrate binding was not observed, possibly because the short sequence and small size of the fold do not provide enough protein structure to support substrate binding.

Although there are many examples of zinc metalloenzymes which have been studied through alteration of their zinc coordination spheres, none have been more extensively reengineered than CA. Because this enzyme has been systematically studied and represents a desirable goal for achieving efficient hydrolytic designed zinc metalloproteins (and is the model system for the designs I will discuss in Chapters 2-5), I

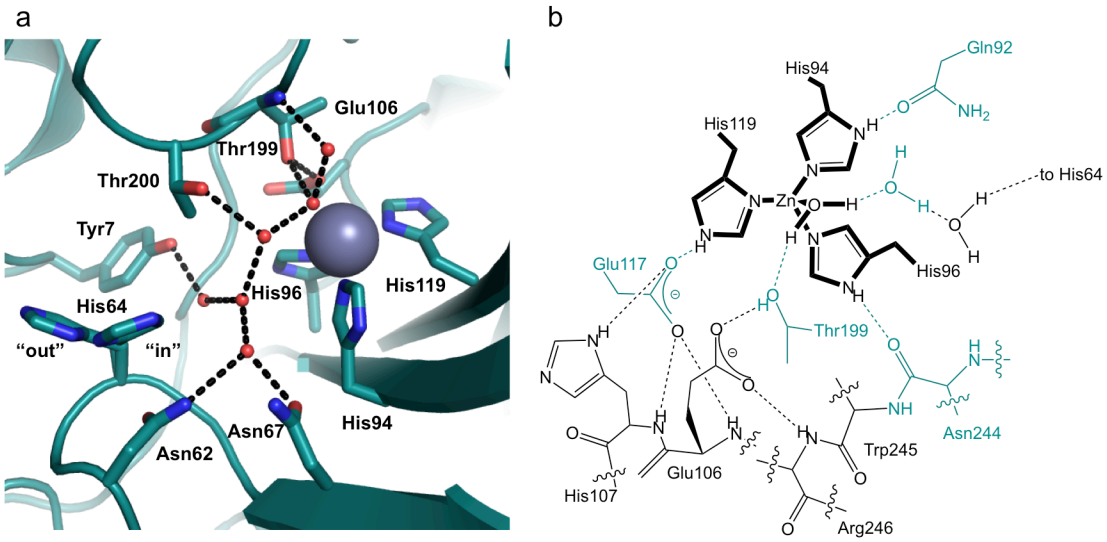

Figure 1-28. Structure of the active site of carbonic anhydrase II. a) The extended active site of CAII showing the ordered water network (pdb 3KS3).[269] Zn(II) is represented by a gray sphere, and the oxygen atoms of water molecules are shown as smaller red spheres. The dotted lines represent presumed hydrogen bonds. Selected amino acids of the active site are shown as stick figures with both the inward and outward orientations of His64 (proton shuttle) shown. This figure was created with PyMOL.[264] b) ChemDraw scheme of the CAII primary and secondary Zn(II)-binding site.

will summarize selected reengineering studies for CA (Figure 1-28). The direct metal-binding residues around the Zn(II) site (His94, His96, and His119) have been replaced by a variety of amino acids (Asp, Glu, Asn, Gln, Cys, Ala) in order to study the effects on the first coordination sphere.[189,251] Early work involved the preparation of the mutant His94Asp, which resulted in a tetrahedral His_2Asp coordination environment with a solvent molecule.[252] A loss in metal-binding affinity of ~10^4-fold was observed (K_d = 15 nM, Table 1-6), probably in part due to movement of the Zn(II) ion ~1 Å towards the Asp residue, resulting in interruption of the evolved His_3 ligand arrangement. Changes in the electrostatics of the site (neutral to anionic substitution) also lead to an increase in the pK_a from 6.8 to > 9.6. A larger loss in Zn(II)-binding affinity is observed when His94 is replaced by Ala (K_d = 270 nM).[251] Substitution of the same His residue with a Cys also results in a loss of Zn(II)-binding affinity of ~10^4-fold (Zn(II) movement is essentially the same as for H94D) and an increase of the pK_a to > 9.5.[251,253] Overall, changes to the charge of the Zn(II) site by addition of charged residues results in severe losses to

catalytic activity (altered Lewis acidity of Zn(II)). However, when alternate neutral residues (Asn, Gln) are introduced, reasonable catalytic activity is retained, although losses in binding affinity of ~10^4-10^5-fold are still observed.[166] Altering the other His positions in similar ways gives similar results.[251,254] Interestingly, a route towards improving the affinity of Zn(II) for CA has also been taken by substituting a nearby residue (Thr199, which forms a stabilizing hydrogen bond with Zn(II)-bound solvent) with Cys, Asp, Glu, or His. Substitution with Cys results in an improvement of the K_d (to 1.1 pM from 4 pM) and some loss in activity (~10^3-fold) due to an alternate conformer with a Zn(II)-hydroxide species.[196,255] In the case of T199E, the affinity is greatly improved (K_d = 20 fM) and the activity is abolished due to displacement of Zn(II)-bound hydroxide.[256] For T199H, the affinity actually decreases 20-fold as the fourth His does not bind to Zn(II).[256] Another example which alters the secondary coordination sphere is mutating T199 to Ala (abolishing the hydrogen bond), which retains the His_3 coordination environment, but results in a 100-fold loss in activity and increase in the pK_a of ~1.5 units. There are also hydrogen-bonding interactions with the unbound nitrogen of each of the His residues in the primary coordination sphere (His94 to Gln92, His119 to Glu117, and His96 to backbone of Asn244). The effects of mutating Gln92 and Glu117 on the pK_a are within ±1 unit with only subtle effects on the K_d, and usually losses in activity of about 10-fold or less. This is likely because the direct metal ligands will often replace the lost hydrogen-bonding partners with other protein sites or even solvent.[257–259] One case in which drastic effects were observed is for the E117Q mutant, in which a 55000-fold loss in activity, increase in the pK_a to > 9, and 1100-fold weaker Zn(II)-binding affinity was reported.[260] These results have been proposed to arise from reversal of the hydrogen bond between residue 117 and His119, stabilizing H119 as a histidinate anion. These studies on the primary and secondary coordination sphere of CA clearly demonstrate many of the characteristics of successful Zn(II) sites that should be taken into account in the protein design approach. Mutagenesis studies demonstrate that neutral His ligands in the primary coordination sphere are essential for enhancing the net positive charge at the Zn(II) site and secondary hydrogen-bonding interactions fine-tune Zn(II)-binding affinities and Zn(II)-OH_2 pK_a's.

Table 1-6. Active site properties for carbonic anhydrase II and selected mutants.

Variant	Coordination Site	K_d (Zn(II)) (pM)	pK_a	CO_2 hydration k_{cat}/K_M (μM^{-1} s^{-1})	*p*NPA hydrolysis (M^{-1} s^{-1})
wild type	His_3O	0.8 ± 0.1[a]	6.8 ± 0.1[b]	110 ± 10[c]	2600 ± 50[c]
Primary coordination sphere mutants					
H94D[d,e]	His_2Asp	15000 ± 5000	≥ 9.6	0.11 ± 0.01	365 ± 70
H94A[e]	His_2	270000 ± 50000		0.012 ± 0.0002	18
H94C[e,f]	His_2Cys	33000 ± 7000	≥ 9.5	0.11 ± 0.01	117 ± 20
Secondary coordination sphere mutants					
T199C[g,h]	His_3Cys	1.1 ± 0.2		0.11 ± 0.02	
T199D[i]	His_3Asp	4 ± 0.02		0.04 ± 0.005	
T199E[i]	His_3Glu	0.02 ± 0.01		0.04 ± 0.005	
T199H[i]	His_3O	77 ± 13	≥ 9.0	0.024 ± 0.002	
T199A	His_3O	60 ± 10[j]	8.3 ± 0.1[c]	1.1 ± 0.05[c]	44 ± 2[c] or 15[k]
E117Q[l]	His_3O	4400 ± 400	≥ 9.0	0.002	3

a. Taken from reference 270.
b. Taken from reference 271.
c. Taken from reference 272.
d. Taken from reference 252.
e. Taken from reference 251.
f. Taken from reference 253.
g. Taken from reference 255.
h. Taken from reference 196.
i. Taken from reference 256.
j. Taken from reference 258.
k. Taken from reference 273.
l. Taken from reference 260.

There are also several reports on the redesign of CA for improved activity towards alternate ester substrates, in which controlling the steric interactions within the hydrophobic pocket can lead to altered substrate specificities and esterase activities.[172,261–263] Therefore, protein design efforts should consider how size and charge interactions of the desired substrates could affect activity and control selectivity.

An example of more extensive reengineering of a zinc metalloenzyme is the computational redesign of the dizinc-containing glyoxalase II enzyme (which hydrolyzes thioester bonds) into a β-lactamase through extensive modifications made through insertion, deletion, and substitution of several active site loops and subsequent point mutations.[33] This redesign altered the metal-binding geometry as well as the substrate-

binding pocket, yet both glyoxalase II and metallo β-lactamase enzymes contain binuclear metal ions essential to the hydrolysis reaction. While a designed dinuclear zinc site is present in the *de novo* designed due ferri (DF) family of proteins (Section 3a)[15], no activity has been reported for any dinuclear zinc site introduced into any *de novo* or preexisting metal scaffold that did not already contain a Zn_2 site.

c) From zinc-binding proteins to hydrolytic metalloenzymes

Until recently, designed zinc metalloproteins have often shown very little, if any, catalytic activity relative to their natural counterparts and few have been structurally characterized. As discussed above, there are examples where preexisting Zn(II) sites have been redesigned for different functions, such as converting gloyoxalase II to a metallo-β-lactamase[33] and converting ZF peptides (closed coordination spheres) to catalytic sites by opening up the first coordination sphere,[236] but generating hydrolytic activity from a "new" zinc site either in a *de novo* or pre-existing protein scaffold had not yet been achieved. In early 2012, I reported the first example of a hydrolytically active *de novo* designed zinc site (Chapter 2).[70] Soon after, Khare *et al.* reported the *de novo* computational redesign and subsequent directed evolution of a Zn(II)-containing mouse adenosine deaminase for efficiently catalyzing the hydrolysis of a model organophosphate substrate.[19] Not long after that, Der *et al.* reported the *de novo* design of a $ZnHis_3$ site at the computationally-designed interface between two copies of the Rab4-binding domain of Rabenosyn which effectively catalyzes the hydrolysis of both *p*NPA and *p*-nitrophenyl phosphate (*p*NPP).[18,39] In the following chapters, I will present my work on the design of a dual-site combined structural and functional metalloprotein for Zn(II)-dependent hydrolysis.

In the second chapter of my thesis, I have described the preparation and characterization of a *de novo* designed protein (a homomeric parallel 3SCC and derivative of the **TRI** peptide sequence), which can bind two different metals in spatially separated sites, with different functions. The sequence used, **TRI**L9CL23H, contains a Cys substitution in the 9th position and His in the 23rd position. The Cys_3 site exploits the Hg(II)-binding studies described in Section 2b in order to function as a structurally

stabilizing metal site. The His_3 site binds transition metals such as Zn(II) and represents a catalytic metal site. The minimal ZnN_3O site achieved in this α-helical coiled coil peptide allows me to address the question of whether a minimal structural unit in a *de novo* protein can achieve significant enzyme-like hydrolytic activity. Further, the inclusion of a structurally stabilizing site sets the stage for comparison to an analogous complex lacking the site and allows for determination of its effects on the activity of the ZnN_3O site. This is particularly important for developing multi-metal site proteins where one may want to incorporate multiple metal sites for a variety of functions.

In the third chapter of my thesis, I address the effects of varying the position of the ZnN_3O site along the sequence of the **TRI** peptide and its insertion into the analogous **Grand** peptide. As discussed in Section 4b, the location of a $ZnHis_2Cys_2$ site in Trx can significantly affect Zn(II)-binding affinities and be correlated with the stability of the complex.[212] One example in which the effects of site location on catalysis was examined also involved the redesign of Trx (in order to incorporate a mononuclear non-heme iron site for superoxide dismutase function), and significant functional differences were observed.[30] By testing two additional site locations along the **TRI** sequence (**TRI**L9HL23C and **TRI**L9CL19H) and its insertion into **Grand** (**GR**L2WL16CL30H), I address whether location of a minimal active site may affect catalytic rate, binding affinity, substrate, solvent, and inhibitor access, Zn(II)-OH_2 pK_a, and catalytic efficiency. Understanding whether these properties may be controlled is important to the design of secondary interactions, including hydrogen-bonding networks.

In the fourth chapter of my thesis, I describe the preparation of additional designed peptides towards the development of hydrogen-bonding networks. I have inserted polar residues (Thr, Asp) into **TRI** and **Grand** sequences around the ZnN_3O site and establish first: whether the peptide can accommodate the substitution and remain folded; second: the binding strength of Zn(II) to each site and how it compares to original sites; and third: the effects these substitutions have on catalytic activity. Then, in order to establish a method for readily interpreting these results in solution, I describe the binding of Cd(II) and Co(II) to **TRI**L2WL23H and Co(II) to those sequences containing polar residues. These metals are commonly used as spectroscopic probes for the Zn(II) site in many native metalloenzymes, including CA.

In the fifth chapter of my thesis, I describe the hydrolytic activity of the designed ZnN_3O site in several of these peptides towards the physiological substrate of CA, CO_2. Although *p*NPA represents an excellent model substrate for measuring the activity of many model complexes, designed proteins, and even native enzymes, it does not represent the evolved activity of the target enzyme, CA. Here, I describe this activity and present comparisons between active sites at different peptide locations and then address how these results may compare to those obtained for *p*NPA hydrolysis.

Finally, in the sixth chapter of my thesis, I conclude with the key achievements from each of my data chapters and propose future directions for this project.

References

(1) Holm, R. H.; Kennepohl, P.; Solomon, E. I. *Chem. Rev.* **1996**, *96*, 2239–2314.
(2) Ragsdale, S. W. *Chem. Rev.* **2006**, *106*, 3317–3337.
(3) Lu, Y.; Berry, S. M.; Pfister, T. D. *Chem. Rev.* **2001**, *101*, 3047–3080.
(4) Barker, P. D. *Curr. Opin. Struct. Biol.* **2003**, *13*, 490–499.
(5) Lu, Y. *Inorg. Chem.* **2006**, *45*, 9930–9940.
(6) Kaplan, J.; DeGrado, W. F. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 11566–11570.
(7) Lu, Y.; Yeung, N.; Sieracki, N.; Marshall, N. M. *Nature* **2009**, *460*, 855–862.
(8) Nanda, V.; Koder, R. L. *Nat. Chem.* **2010**, *2*, 15–24.
(9) Harris, K. L.; Lim, S.; Franklin, S. J. *Inorg. Chem.* **2006**, *45*, 10002–10012.
(10) DeGrado, W. F.; Summa, C. M.; Pavone, V.; Nastri, F.; Lombardi, A. *Annu. Rev. Biochem.* **1999**, *68*, 779–819.
(11) Koder, R. L.; Dutton, P. L. *Dalton Trans.* **2006**, 3045–3051.
(12) DeGrado, W. F.; Raleigh, D. P.; Handel, T. *Curr. Opin. Struct. Biol.* **1991**, *1*, 984–993.
(13) Hong, J.; Kharenko, O. A.; Ogawa, M. Y. *Inorg. Chem.* **2006**, *45*, 9974–9984.
(14) Yeung, N.; Lin, Y.-W.; Gao, Y.-G.; Zhao, X.; Russell, B. S.; Lei, L.; Miner, K. D.; Robinson, H.; Lu, Y. *Nature* **2009**, *462*, 1079–1082.
(15) 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.
(16) Miner, K. D.; Mukherjee, A.; Gao, Y.-G.; Null, E. L.; Petrik, I. D.; Zhao, X.; Yeung, N.; Robinson, H.; Lu, Y. *Angew. Chem., Int. Ed.* **2012**, *51*, 5589–5592.
(17) Liu, X.; Yu, Y.; Hu, C.; Zhang, W.; Lu, Y.; Wang, J. *Angew. Chem., Int. Ed.* **2012**, *51*, 4312–4316.
(18) Der, B. S.; Edwards, D. R.; Kuhlman, B. *Biochemistry* **2012**, *51*, 3933–3940.
(19) Khare, S. D.; Kipnis, Y.; Greisen, P. J.; Takeuchi, R.; Ashani, Y.; Goldsmith, M.; Song, Y.; Gallaher, J. L.; Silman, I.; Leader, H.; Sussman, J. L.; Stoddard, B. L.; Tawfik, D. S.; Baker, D. *Nat. Chem. Biol.* **2012**, *8*, 294–300.
(20) Wong-Deyrup, S. W.; Prasannan, C.; Dupureur, C. M.; Franklin, S. J. *J. Biol. Inorg. Chem.* **2012**, *17*, 387–398.
(21) Garner, D. K.; Liang, L.; Barrios, D. A.; Zhang, J.; Lu, Y. *ACS Catal.* **2011**, *1*, 1083–1089.
(22) Mayer, C.; Gillingham, D. G.; Ward, T. R.; Hilvert, D. *Chem. Commun. (Cambridge, U. K.)* **2011**, *47*, 12068–12070.
(23) Allard, M.; Dupont, C.; Muñoz Robles, V.; Doucet, N.; Lledós, A.; Maréchal, J.-D.; Urvoas, A.; Mahy, J.-P.; Ricoux, R. *ChemBioChem* **2012**, *13*, 240–251.
(24) Bos, J.; Fusetti, F.; Driessen, A. J. M.; Roelfes, G. *Angew. Chem., Int. Ed.* **2012**, *51*, 7472–7475.
(25) Matsuo, T.; Imai, C.; Yoshida, T.; Saito, T.; Hayashi, T.; Hirota, S. *Chem. Commun. (Cambridge, U. K.)* **2012**, *48*, 1662–1664.
(26) Ringenberg, M. R.; Ward, T. R. *Chem. Commun. (Cambridge, U. K.)* **2011**, *47*, 8470–8476.
(27) Steinreiber, J.; Ward, T. R. *Coord. Chem. Rev.* **2008**, *252*, 751–766.
(28) Ward, T. R. *Acc. Chem. Res.* **2011**, *44*, 47–57.

(29) Hellinga, H. W. *Fold. Des.* **1998**, *3*, R1–R8.
(30) Benson, D. E.; Wisz, M. S.; Hellinga, H. W. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 6292–6297.
(31) DeGrado, W. F.; Wasserman, Z. R. *Science* **1989**, *243*, 622–628.
(32) 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.
(33) Park, H.-S.; Nam, S.; Lee, J. K.; Yoon, C. N.; Mannervik, B.; Benkovic, S. J.; Kim, H.-S. *Science* **2006**, *311*, 535–538.
(34) Clark, K. M.; Yu, Y.; Marshall, N. M.; Sieracki, N. A.; Nilges, M. J.; Blackburn, N. J.; van der Donk, W. A.; Lu, Y. *J. Am. Chem. Soc.* **2010**, *132*, 10093–10101.
(35) Fujieda, N.; Hasegawa, A.; Ishihama, K.-I.; Itoh, S. *Chem.--Asian J.* **2012**, *7*, 1203–1207.
(36) Salgado, E. N.; Radford, R. J.; Tezcan, F. A. *Acc. Chem. Res.* **2010**, *43*, 661–672.
(37) Radford, R. J.; Brodin, J. D.; Salgado, E. N.; Tezcan, F. A. *Coord. Chem. Rev.* **2011**, *255*, 790–803.
(38) Brodin, J. D.; Ambroggio, X. I.; Tang, C.; Parent, K. N.; Baker, T. S.; Tezcan, F. A. *Nat. Chem.* **2012**, *4*, 375–382.
(39) Der, B. S.; Machius, M.; Miley, M. J.; Mills, J. L.; Szyperski, T.; Kuhlman, B. *J. Am. Chem. Soc.* **2012**, *134*, 375–385.
(40) Anderson, J. L. R.; Koder, R. L.; Moser, C. C.; Dutton, P. L. *Biochem. Soc. Trans.* **2008**, *36*, 1106–1111.
(41) Hodges, R. S. *Biochem. Cell Biol.* **1996**, *74*, 133–154.
(42) Apostolovic, B.; Danial, M.; Klok, H.-A. *Chem. Soc. Rev.* **2010**, *39*, 3541–3575.
(43) Woolfson, D. N. *Adv. Protein Chem.* **2005**, *70*, 79–112.
(44) Hodges, R. S.; Saund, A. K.; Chong, P. C.; St-Pierre, S. A.; Reid, R. E. *J. Biol. Chem.* **1981**, *256*, 1214–1224.
(45) Crick, F. H. C. *Acta Crystallogr.* **1953**, *6*, 689–697.
(46) DeGrado, W. F.; Lear, J. D. *J. Am. Chem. Soc.* **1985**, *107*, 7684–7689.
(47) Kamtekar, S.; Schiffer, J.; Xiong, H.; Babik, J.; Hecht, M. *Science* **1993**, *262*, 1680–1685.
(48) Lau, S. Y.; Taneja, A. K.; Hodges, R. S. *J. Biol. Chem.* **1984**, *259*, 13253–13261.
(49) Handel, T. T.; Williams, S. A.; DeGrado, W. F. *Science* **1993**, *261*, 879–885.
(50) Lovejoy, B.; Choe, S.; Cascio, D.; McRorie, D.; DeGrado, W.; Eisenberg, D. *Science* **1993**, *259*, 1288–1293.
(51) Ogihara, N. L.; Weiss, M. S.; Degrado, W. F.; Eisenberg, D. *Protein Sci.* **1997**, *6*, 80–88.
(52) Betz, S.; Fairman, R.; O'Neil, K.; Lear, J.; Degrado, W. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **1995**, *348*, 81–88.
(53) Ghosh, D.; Pecoraro, V. L. *Inorg. Chem.* **2004**, *43*, 7902–7915.
(54) Kharenko, O. A.; Ogawa, M. Y. *J. Inorg. Biochem.* **2004**, *98*, 1971–1974.
(55) Shiga, D.; Nakane, D.; Inomata, T.; Funahashi, Y.; Masuda, H.; Kikuchi, A.; Oda, M.; Noda, M.; Uchiyama, S.; Fukui, K.; Kanaori, K.; Tajima, K.; Takano, Y.; Nakamura, H.; Tanaka, T. *J. Am. Chem. Soc.* **2010**, *132*, 18191–18198.
(56) Shiga, D.; Nakane, D.; Inomata, T.; Masuda, H.; Oda, M.; Noda, M.; Uchiyama, S.; Fukui, K.; Takano, Y.; Nakamura, H.; Mizuno, T.; Tanaka, T. *Biopolymers* **2009**, *91*, 907–916.

(57) Shiga, D.; Funahashi, Y.; Masuda, H.; Kikuchi, A.; Noda, M.; Uchiyama, S.; Fukui, K.; Kanaori, K.; Tajima, K.; Takano, Y.; Nakamura, H.; Kamei, M.; Tanaka, T. *Biochemistry* **2012**, *51*, 7901–7907.
(58) Xie, F.; Sutherland, D. E. K.; Stillman, M. J.; Ogawa, M. Y. *J. Inorg. Biochem.* **2010**, *104*, 261–267.
(59) Peacock, A. F. A.; Iranzo, O.; Pecoraro, V. L. *Dalton Trans.* **2009**, *9226*, 2271–2280.
(60) Iranzo, O.; Chakraborty, S.; Hemmingsen, L.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2011**, *133*, 239–251.
(61) Matzapetakis, M.; Ghosh, D.; Weng, T.-C.; Penner-Hahn, J. E.; Pecoraro, V. L. *J. Biol. Inorg. Chem.* **2006**, *11*, 876–890.
(62) Neupane, K. P.; Pecoraro, V. L. *Angew. Chem., Int. Ed.* **2010**, *49*, 8177–8180.
(63) Pecoraro, V. L.; Peacock, A. F. A.; Iranzo, O.; Luczkowski, M. *ACS Symp. Ser.* **2009**, *1012*, 183–197.
(64) Ghosh, D.; Pecoraro, V. L. *Curr. Opin. Chem. Biol.* **2005**, *9*, 97–103.
(65) Iranzo, O.; Ghosh, D.; Pecoraro, V. L. *Inorg. Chem.* **2006**, *45*, 9959–9973.
(66) Iranzo, O.; Jakusch, T.; Lee, K.-H.; Hemmingsen, L.; Pecoraro, V. L. *Chem.--Eur. J.* **2009**, *15*, 3761–3772.
(67) Chakraborty, S.; Touw, D. S.; Peacock, A. F. A.; Stuckey, J.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2010**, *132*, 13240–13250.
(68) Touw, D. S.; Nordman, C. E.; Stuckey, J. A.; Pecoraro, V. L. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 11969–11974.
(69) Zaykov, A. N.; MacKenzie, K. R.; Ball, Z. T. *Chem.--Eur. J.* **2009**, *15*, 8961–8965.
(70) Zastrow, M. L.; Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. *Nat. Chem.* **2012**, *4*, 118–123.
(71) Dieckmann, G. R.; McRorie, D. K.; Lear, J. D.; Sharp, K. A.; DeGrado, W. F.; Pecoraro, V. L. *J. Mol. Biol.* **1998**, *280*, 897–912.
(72) 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.
(73) Bryson, J. W.; Desjarlais, J. R.; Handel, T. M.; DeGrado, W. F. *Protein Sci.* **1998**, *7*, 1404–1414.
(74) Walsh, S. T. R.; Cheng, H.; Bryson, J. W.; Roder, H.; DeGrado, W. F. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 5486–5491.
(75) Regan, L.; DeGrado, W. F. *Science* **1988**, *241*, 976–978.
(76) Handel, T.; DeGrado, W. F. *J. Am. Chem. Soc.* **1990**, *112*, 6710–6711.
(77) Hill, R. B.; Raleigh, D. P.; Lombardi, A.; DeGrado, W. F. *Acc. Chem. Res.* **2000**, *33*, 745–754.
(78) Schnepf, R.; Hörth, P.; Bill, E.; Wieghardt, K.; Hildebrandt, P.; Haehnel, W. *J. Am. Chem. Soc.* **2001**, *123*, 2186–2195.
(79) Chakraborty, S.; Kravitz, J. Y.; Thulstrup, P. W.; Hemmingsen, L.; DeGrado, W. F.; Pecoraro, V. L. *Angew. Chem., Int. Ed.* **2011**, *50*, 2049–2053.
(80) Regan, L.; Clarke, N. D. *Biochemistry* **1990**, *29*, 10878–10883.

(81) Reig, A. J.; Pires, M. M.; Snyder, R. A.; Wu, Y.; Jo, H.; Kulp, D. W.; Butch, S. E.; Calhoun, J. R.; Szyperski, T.; Szyperski, T. G.; Solomon, E. I.; DeGrado, W. F. *Nat. Chem.* **2012**, *4*, 900–906.
(82) Faiella, M.; Andreozzi, C.; de Rosales, R. T. M.; Pavone, V.; Maglio, O.; Nastri, F.; DeGrado, W. F.; Lombardi, A. *Nat. Chem. Biol.* **2009**, *5*, 882–884.
(83) Maglio, O.; Nastri, F.; Martin de Rosales, R. T.; Faiella, M.; Pavone, V.; DeGrado, W. F.; Lombardi, A. *C. R. Chim.* **2007**, *10*, 703–720.
(84) Calhoun, J. R.; Nastri, F.; Maglio, O.; Pavone, V.; Lombardi, A.; DeGrado, W. F. *Biopolymers* **2005**, *80*, 264–278.
(85) Cheng, R. P.; Gellman, S. H.; DeGrado, W. F. *Chem. Rev.* **2001**, *101*, 3219–3232.
(86) Lacroix, E.; Kortemme, T. *Curr. Opin. Struct. Biol.* **1999**, *9*, 487–493.
(87) Schneider, J. P.; Kelly, J. W. *J. Am. Chem. Soc.* **1995**, *117*, 2533–2546.
(88) Platt, G.; Searle, M. S.; Chung, C.-W. *Chem. Commun. (Cambridge, U. K.)* **2001**, 1162–1163.
(89) Venkatraman, J.; Naganagowda, G. A.; Sudha, R.; Balaram, P. *Chem. Commun. (Cambridge, U. K.)* **2001**, 2660–2661.
(90) Pessi, A.; Bianchi, E.; Crameri, A.; Venturini, S.; Tramontano, A.; Sollazzo, M. *Nature* **1993**, *362*, 367–369.
(91) Zhu, C.; Zhang, C.; Liang, H.; Lai, L. *Protein Cell* **2011**, *2*, 1006–1013.
(92) Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D. *Science* **2003**, *302*, 1364–1368.
(93) Struthers, M. D.; Cheng, R. P.; Imperiali, B. *J. Am. Chem. Soc.* **1996**, *118*, 3073–3081.
(94) Struthers, M. D.; Cheng, R. P.; Imperiali, B. *Science* **1996**, *271*, 342–345.
(95) Bonomo, R. P.; Casella, L.; De Gioia, L.; Molinari, H.; Impellizzeri, G.; Jordan, T.; Pappalardo, G.; Purrello, R.; Rizzarelli, E. *Dalton Trans.* **1997**, 2387–2390.
(96) Butterfield, S. M.; Cooper, W. J.; Waters, M. L. *J. Am. Chem. Soc.* **2005**, *127*, 24–25.
(97) Butterfield, S. M.; Goodman, C. M.; Rotello, V. M.; Waters, M. L. *Angew. Chem., Int. Ed.* **2004**, *43*, 724–727.
(98) Hughes, R. M.; Waters, M. L. *Curr. Opin. Struct. Biol.* **2006**, *16*, 514–524.
(99) Cheng, R. P.; Fisher, S. L.; Imperiali, B. *J. Am. Chem. Soc.* **1996**, *118*, 11349–11356.
(100) Daugherty, R. G.; Wasowicz, T.; Gibney, B. R.; DeRose, V. J. *Inorg. Chem.* **2002**, *41*, 2623–2632.
(101) Kozísek, M.; Svatos, A.; Budesínský, M.; Muck, A.; Bauer, M. C.; Kotrba, P.; Ruml, T.; Havlas, Z.; Linse, S.; Rulísek, L. *Chem.--Eur. J.* **2008**, *14*, 7836–7846.
(102) 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.
(103) Turano, P. *Inorg. Chem.* **2004**, *43*, 7945–7952.
(104) Maret, W.; Li, Y. *Chem. Rev.* **2009**, *109*, 4682–4707.
(105) Finn, B. E.; Drakenberg, T. *Adv. Inorg. Chem.* **1999**, *46*, 441–494.
(106) Doerr, A. J.; McLendon, G. L. *Inorg. Chem.* **2004**, *43*, 7916–7925.
(107) Ghadiri, M. R.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, *114*, 825–831.
(108) Tsurkan, M. V; Ogawa, M. Y. *Inorg. Chem.* **2007**, *46*, 6849–6851.
(109) Zaykov, A. N.; Popp, B. V; Ball, Z. T. *Chem.--Eur. J.* **2010**, *16*, 6651–6659.

(110) Suzuki, K.; Hiroaki, H.; Kohda, D.; Nakamura, H.; Tanaka, T. *J. Am. Chem. Soc.* **1998**, *120*, 13008–13015.
(111) Li, X.; Suzuki, K.; Kanaori, K.; Tajima, K.; Kashiwada, A.; Hiroaki, H.; Kohda, D.; Tanaka, T. *Protein Sci.* **2000**, *9*, 1327–1333.
(112) Ghadiri, M. R.; Choi, C. *J. Am. Chem. Soc.* **1990**, *112*, 1630–1632.
(113) Ghadiri, M. R.; Fernholz, A. K. *J. Am. Chem. Soc.* **1990**, *112*, 9633–9635.
(114) Ghadiri, M. R.; Case, M. A. *Angew. Chem., Int. Ed.* **1993**, *32*, 1594–1597.
(115) Ghadiri, M. R.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, *114*, 4000–4002.
(116) Kharenko, O. A.; Kennedy, D. C.; Demeler, B.; Maroney, M. J.; Ogawa, M. Y. *J. Am. Chem. Soc.* **2005**, *127*, 7678–7679.
(117) Kiyokawa, T.; Kanaori, K.; Tajima, K.; Koike, M.; Mizuno, T.; Oku, J.-I.; Tanaka, T. *J. Pept. Res.* **2004**, *63*, 347–353.
(118) Murase, S.; Ishino, S.; Ishino, Y.; Tanaka, T. *J. Biol. Inorg. Chem.* **2012**, *17*, 791–799.
(119) O'Neil, K. T.; DeGrado, W. F. *Science* **1990**, *250*, 646–651.
(120) Iranzo, O.; Cabello, C.; Pecoraro, V. L. *Angew. Chem., Int. Ed.* **2007**, *46*, 6688–6691.
(121) Lee, K.-H.; Cabello, C.; Hemmingsen, L.; Marsh, E. N. G.; Pecoraro, V. L. *Angew. Chem., Int. Ed.* **2006**, *45*, 2864–2868.
(122) Peacock, A. F. A.; Hemmingsen, L.; Pecoraro, V. L. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 16566–16571.
(123) Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. *Angew. Chem., Int. Ed.* **2009**, *48*, 7371–7374.
(124) Farrer, B. T.; McClure, C. P.; Penner-Hahn, J. E.; Pecoraro, V. L. *Inorg. Chem.* **2000**, *39*, 5422–5423.
(125) Farrer, B. T.; Harris, N. P.; Balchus, K. E.; Pecoraro, V. L. *Biochemistry* **2001**, *40*, 14696–14705.
(126) Farrer, B. T.; Pecoraro, V. L. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3760–3765.
(127) Matzapetakis, M.; Farrer, B. T.; Weng, T.-C.; Hemmingsen, L.; Penner-Hahn, J. E.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2002**, *124*, 8042–8054.
(128) Matzapetakis, M.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2005**, *127*, 18229–18233.
(129) Chakraborty, S.; Iranzo, O.; Zuiderweg, E. R. P.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2012**, *134*, 6191–6203.
(130) Rulíšek, L.; Vondrášek, J. *J. Inorg. Biochem.* **1998**, *71*, 115–127.
(131) Zheng, H.; Chruszcz, M.; Lasota, P.; Lebioda, L.; Minor, W. *J. Inorg. Biochem.* **2008**, *102*, 1765–1776.
(132) Wright, J. G.; Tsang, H. T.; Penner-Hahn, J. E.; O'Halloran, T. V. *J. Am. Chem. Soc.* **1990**, *112*, 2434–2435.
(133) Faller, P.; Ctortecka, B.; Tröger, W.; Butz, T.; Vasák, M. *J. Biol. Inorg. Chem.* **2000**, *5*, 393–401.
(134) Łuczkowski, M.; Stachura, M.; Schirf, V.; Demeler, B.; Hemmingsen, L.; Pecoraro, V. L. *Inorg. Chem.* **2008**, *47*, 10875–10888.
(135) Geremia, S.; Di Costanzo, L.; Randaccio, L.; Engel, D. E.; Lombardi, A.; Nastri, F.; DeGrado, W. F. *J. Am. Chem. Soc.* **2005**, *127*, 17266–17276.
(136) Maglio, O.; Nastri, F.; Pavone, V.; Lombardi, A.; DeGrado, W. F. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3772–3777.

(137) DeGrado, W. F.; Di Costanzo, L.; Geremia, S.; Lombardi, A.; Pavone, V.; Randaccio, L. *Angew. Chem., Int. Ed.* **2003**, *42*, 417–420.
(138) Di Costanzo, L.; Wade, H.; Geremia, S.; Randaccio, L.; Pavone, V.; DeGrado, W. F.; Lombardi, A. *J. Am. Chem. Soc.* **2001**, *123*, 12749–12757.
(139) Pasternak, A.; Kaplan, J.; Lear, J. D.; Degrado, W. F. *Protein Sci.* **2001**, *10*, 958–969.
(140) Maglio, O.; Nastri, F.; Calhoun, J. R.; Lahr, S.; Wade, H.; Pavone, V.; DeGrado, W. F.; Lombardi, A. *J. Biol. Inorg. Chem.* **2005**, *10*, 539–549.
(141) Lahr, S. J.; Engel, D. E.; Stayrook, S. E.; Maglio, O.; North, B.; Geremia, S.; Lombardi, A.; DeGrado, W. F. *J. Mol. Biol.* **2005**, *346*, 1441–1454.
(142) Wade, H.; Stayrook, S. E.; Degrado, W. F. *Angew. Chem., Int. Ed.* **2006**, *45*, 4951–4954.
(143) Wei, P.-P.; Skulan, A. J.; Wade, H.; DeGrado, W. F.; Solomon, E. I. *J. Am. Chem. Soc.* **2005**, *127*, 16098–16106.
(144) Summa, C. M.; Rosenblatt, M. M.; Hong, J.-K.; Lear, J. D.; DeGrado, W. F. *J. Mol. Biol.* **2002**, *321*, 923–938.
(145) Marsh, E. N. G.; DeGrado, W. F. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5150–5154.
(146) Calhoun, J. R.; Kono, H.; Lahr, S.; Wang, W.; DeGrado, W. F.; Saven, J. G. *J. Mol. Biol.* **2003**, *334*, 1101–1115.
(147) Calhoun, J. R.; Liu, W.; Spiegel, K.; Dal Peraro, M.; Klein, M. L.; Valentine, K. G.; Wand, A. J.; DeGrado, W. F. *Structure* **2008**, *16*, 210–215.
(148) Calhoun, J. R.; Bell, C. B.; Smith, T. J.; Thamann, T. J.; DeGrado, W. F.; Solomon, E. I. *J. Am. Chem. Soc.* **2008**, *130*, 9188–9189.
(149) Bell, C. B.; Calhoun, J. R.; Bobyr, E.; Wei, P.-P.; Hedman, B.; Hodgson, K. O.; Degrado, W. F.; Solomon, E. I. *Biochemistry* **2009**, *48*, 59–73.
(150) Ball, Z. T. *Acc. Chem. Res.* **2013**, *46*, 560–70.
(151) Litowski, J. R.; Hodges, R. S. *J. Biol. Chem.* **2002**, *277*, 37272–37279.
(152) Sambasivan, R.; Ball, Z. T. *J. Am. Chem. Soc.* **2010**, *132*, 9289–9291.
(153) Popp, B. V; Ball, Z. T. *J. Am. Chem. Soc.* **2010**, *132*, 6660–6662.
(154) Popp, B. V.; Ball, Z. T. *Chem. Sci.* **2011**, *2*, 690–695.
(155) Chen, Z.; Popp, B. V; Bovet, C. L.; Ball, Z. T. *ACS Chem. Biol.* **2011**, *6*, 920–925.
(156) Chen, Z.; Vohidov, F.; Coughlin, J. M.; Stagg, L. J.; Arold, S. T.; Ladbury, J. E.; Ball, Z. T. *J. Am. Chem. Soc.* **2012**, *134*, 10138–10145.
(157) Popp, B. V; Chen, Z.; Ball, Z. T. *Chem. Commun. (Cambridge, U. K.)* **2012**, *48*, 7492–7494.
(158) Andreini, C.; Banci, L.; Bertini, I.; Rosato, A. *J. Proteome Res.* **2006**, *5*, 196–201.
(159) Keilin, D.; Mann, T. *Nature* **1939**, *144*, 442–443.
(160) Vallee, B. L.; Neurath, H. *J. Am. Chem. Soc.* **1954**, *76*, 5006–5007.
(161) Vallee, B. L. *Physiol. Rev.* **1959**, *39*, 443–490.
(162) Vallee, B. L.; Auld, D. S. *Biochemistry* **1990**, *29*, 5647–5659.
(163) Auld, D. S. *Biometals* **2001**, *14*, 271–313.
(164) Gomis-Rüth, F. X.; Stöcker, W.; Huber, R.; Zwilling, R.; Bode, W. *J. Mol. Biol.* **1993**, *229*, 945–968.
(165) Sträter, N.; Klabunde, T.; Tucker, P.; Witzel, H.; Krebs, B. *Science* **1995**, *268*, 1489–1492.

(166) Lesburg, C. A.; Huang, C.; Christianson, D. W.; Fierke, C. A. *Biochemistry* **1997**, *36*, 15780–15791.
(167) McCall, K. A.; Huang, C.; Fierke, C. A. *J. Nutr.* **2000**, *130*, 1437S–1446S.
(168) Sigel, H.; Martin, R. B. *Chem. Soc. Rev.* **1994**, *23*, 83–91.
(169) Vallee, B. L.; Auld, D. S. *Biochemistry* **1993**, *32*, 6493–6500.
(170) Maret, W. *J. Inorg. Biochem.* **2012**, *111*, 110–116.
(171) Verpoorte, J. A.; Mehta, S.; Edsall, J. T. *J. Biol. Chem.* **1967**, *242*, 4221–4229.
(172) Gould, S. M.; Tawfik, D. S. *Biochemistry* **2005**, *44*, 5444–5452.
(173) Wouters, M. A.; Husain, A. *J. Mol. Biol.* **2001**, *314*, 1191–1207.
(174) Supuran, C. T. *Nat. Rev. Drug Discov.* **2008**, *7*, 168–181.
(175) Supuran, C. T.; Di Fiore, A.; Alterio, V.; Monti, S. M.; De Simone, G. *Curr. Pharm. Des.* **2010**, *16*, 3246–3254.
(176) Imtaiyaz Hassan, M.; Shajee, B.; Waheed, A.; Ahmad, F.; Sly, W. S. *Bioorg. Med. Chem.* **2013**, *21*, 1570–1582.
(177) Namuswe, F.; Berg, J. M. *J. Inorg. Biochem.* **2012**, *111*, 146–149.
(178) Håkansson, K.; Carlsson, M.; Svensson, L. A.; Liljas, A. *J. Mol. Biol.* **1992**, *227*, 1192–1204.
(179) Chang, X.; Jorgensen, A. M.; Bardrum, P.; Led, J. J. *Biochemistry* **1997**, *36*, 9409–9422.
(180) Fujinaga, M.; James, M. N. *J. Mol. Biol.* **1987**, *195*, 373–396.
(181) Banerjee, S.; Wei, B.; Bhattacharyya-Pakrasi, M.; Pakrasi, H. B.; Smith, T. J. *J. Mol. Biol.* **2003**, *333*, 1061–1069.
(182) Parkin, G. *Chem. Rev.* **2004**, *104*, 699–767.
(183) Zastrow, M. L.; Pecoraro, V. L. *Coord. Chem. Rev.* **2013**, *257*, 2565–2588.
(184) Yu, F.; Cangelosi, V. M.; Zastrow, M. L.; Tegoni, M. M.; Plegaria, J. S.; Tebo, A. G.; Mocny, C. S.; Ruckthong, L.; Qayuum, H.; Pecoraro, V. L. *Chem. Rev.* **2013**, *submitted.*
(185) Schurer, G.; Clark, T. *Patai's Chemistry of Functional Groups* **2009**, 1–29.
(186) Galdes, A. In *Inorganic Biochemistry*; Hill, H. A. O., Ed.; The Royal Society of Chemistry: London, 1981; pp. 216–248.
(187) Liang, J. Y.; Lipscomb, W. N. *Biochemistry* **1988**, *27*, 8676–8682.
(188) Lu, D.; Voth, G. A. *J. Am. Chem. Soc.* **1998**, *120*, 4006–4014.
(189) Christianson, D. W.; Fierke, C. A. *Acc. Chem. Res.* **1996**, *29*, 331–339.
(190) Christianson, D. W.; Lipscomb, W. N. *Acc. Chem. Res.* **1989**, *22*, 62–69.
(191) Jaffe, E. K. *Bioorg. Chem.* **2004**, *32*, 316–325.
(192) Ryde, U. *Int. J. Quantum Chem.* **1994**, *52*, 1229–1243.
(193) Myers, L. C.; Terranova, M. P.; Ferentz, A. E.; Wagner, G.; Verdine, G. L. *Science* **1993**, *261*, 1164–1167.
(194) Penner-Hahn, J. *Curr. Opin. Chem. Biol.* **2007**, *11*, 166–171.
(195) Hitomi, Y.; Outten, C. E.; O'Halloran, T. V *J. Am. Chem. Soc.* **2001**, *123*, 8614–8615.
(196) Kiefer, L. L.; Krebs, J. F.; Paterno, S. A.; Fierke, C. A. *Biochemistry* **1993**, *32*, 9896–9900.
(197) McCall, K. A.; Fierke, C. A. *Biochemistry* **2004**, *43*, 3979–3986.
(198) Maret, W. *J. Nutr.* **2003**, *133*, 1460s–1462s.
(199) Müller, H. N.; Skerra, A. *Biochemistry* **1994**, *33*, 14126–14135.

(200) Schmidt, A. M.; Müller, H. N.; Skerra, A. *Chem. Biol.* **1996**, *3*, 645–653.
(201) Vita, C.; Roumestand, C.; Toma, F.; Ménez, A. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 6404–6408.
(202) Lindskog, S.; Nyman, P. O. *Biochim. Biophys. Acta* **1964**, *85*, 462–474.
(203) Klemba, M.; Regan, L. *Biochemistry* **1995**, *34*, 10094–10100.
(204) Gockel, P.; Vahrenkamp, H.; Zuberbuhler, A. D. *Helv. Chim. Acta* **1993**, *76*, 511–520.
(205) Vogler, R.; Vahrenkamp, H. *Eur. J. Inorg. Chem.* **2002**, *2002*, 761–766.
(206) Vallee, B. L.; Auld, D. S. *Acc. Chem. Res.* **1993**, *26*, 543–551.
(207) Reddi, A. R.; Guzman, T. R.; Breece, R. M.; Tierney, D. L.; Gibney, B. R. *J. Am. Chem. Soc.* **2007**, *129*, 12815–12827.
(208) Klemba, M.; Gardner, K. H.; Marino, S.; Clarke, N. D.; Regan, L. *Nat. Struct. Biol.* **1995**, *2*, 368–373.
(209) Jacques, A.; Mettra, B.; Lebrun, V.; Latour, J.-M.; Sénèque, O. *Chem.--Eur. J.* **2013**, *19*, 3921–3931.
(210) Negi, S.; Imanishi, M.; Matsumoto, M.; Sugiura, Y. *Chem.--Eur. J.* **2008**, *14*, 3236–3249.
(211) Imperiali, B.; Kapoor, T. M. *Tetrahedron* **1993**, *49*, 3501–3510.
(212) Wisz, M. S.; Garrett, C. Z.; Hellinga, H. W. *Biochemistry* **1998**, *37*, 8269–8277.
(213) Reddi, A. R.; Gibney, B. R. *Biochemistry* **2007**, *46*, 3745–3758.
(214) Petros, A. K.; Reddi, A. R.; Kennedy, M. L.; Hyslop, A. G.; Gibney, B. R. *Inorg. Chem.* **2006**, *45*, 9941–9958.
(215) Hellinga, H. W.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 763–785.
(216) Krizek, B. A.; Merkle, D. L.; Berg, J. M. *Inorg. Chem.* **1993**, *32*, 937–940.
(217) Outten, C. E.; O'Halloran, T. V *Science* **2001**, *292*, 2488–2492.
(218) Matthews, J. M.; Loughlin, F. E.; Mackay, J. P. *Curr. Opin. Struct. Biol.* **2008**, *18*, 484–490.
(219) Salgado, E. N.; Faraone-Mennella, J.; Tezcan, F. A. *J. Am. Chem. Soc.* **2007**, *129*, 13374–13375.
(220) Salgado, E. N.; Lewis, R. A.; Mossin, S.; Rheingold, A. L.; Tezcan, F. A. *Inorg. Chem.* **2009**, *48*, 2726–2728.
(221) Liu, Y.; Kuhlman, B. *Nucleic Acids Res.* **2006**, *34*, W235–W238.
(222) Salgado, E. N.; Ambroggio, X. I.; Brodin, J. D.; Lewis, R. A.; Kuhlman, B.; Tezcan, F. A. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 1827–1832.
(223) Radford, R. J.; Lawrenz, M.; Nguyen, P. C.; McCammon, J. A.; Tezcan, F. A. *Chem. Commun. (Cambridge, U. K.)* **2011**, *47*, 313–315.
(224) Nomura, A.; Sugiura, Y. *Inorg. Chem.* **2004**, *43*, 1708–1713.
(225) Dhanasekaran, M.; Negi, S.; Sugiura, Y. *Acc. Chem. Res.* **2006**, *39*, 45–52.
(226) Higaki, J. N.; Haymore, B. L.; Chen, S.; Fletterick, R. J.; Craik, C. S. *Biochemistry* **1990**, *29*, 8582–8586.
(227) Halfon, S.; Craik, C. S. *J. Am. Chem. Soc.* **1996**, *118*, 1227–1228.
(228) Willett, W. S.; Gillmor, S. A.; Perona, J. J.; Fletterick, R. J.; Craik, C. S. *Biochemistry* **1995**, *34*, 2172–2180.
(229) Willett, W. S.; Brinen, L. S.; Fletterick, R. J.; Craik, C. S. *Biochemistry* **1996**, *35*, 5992–5998.

(230) Brinen, L. S.; Willett, W. S.; Craik, C. S.; Fletterick, R. J. *Biochemistry* **1996**, *35*, 5999–6009.
(231) Jensen, K. K.; Martini, L.; Schwartz, T. W. *Biochemistry* **2001**, *40*, 938–945.
(232) Evers, T. H.; Appelhof, M. A. M.; de Graaf-Heuvelmans, P. T. H. M.; Meijer, E. W.; Merkx, M. *J. Mol. Biol.* **2007**, *374*, 411–425.
(233) Mizuno, T.; Murao, K.; Tanabe, Y.; Oda, M.; Tanaka, T. *J. Am. Chem. Soc.* **2007**, *129*, 11378–11383.
(234) Liu, H.; Schmidt, J. J.; Bachand, G. D.; Rizk, S. S.; Looger, L. L.; Hellinga, H. W.; Montemagno, C. D. *Nat. Mater.* **2002**, *1*, 173–177.
(235) Dwyer, M. A.; Looger, L. L.; Hellinga, H. W. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11255–11260.
(236) Merkle, D. L.; Schmidt, M. H.; Berg, J. M. *J. Am. Chem. Soc.* **1991**, *113*, 5450–5451.
(237) Nomura, A.; Sugiura, Y. *Inorg. Chem.* **2002**, *41*, 3693–3698.
(238) Negi, S.; Itazu, M.; Imanishi, M.; Nomura, A.; Sugiura, Y. *Biochem. Biophys. Res. Commun.* **2004**, *325*, 421–425.
(239) Itoh, T.; Fujii, Y.; Tada, T.; Yoshikawa, Y.; Hisada, H. *Bull. Chem. Soc. Jpn.* **1996**, *69*, 1265–1274.
(240) King, D. A.; Zhang, L.; Guarente, L.; Marmorstein, R. *Nat. Struct. Biol.* **1999**, *6*, 64–71.
(241) Hori, Y.; Suzuki, K.; Okuno, Y.; Nagaoka, M.; Futaki, S.; Sugiura, Y. *J. Am. Chem. Soc.* **2000**, *122*, 7648–7653.
(242) Salgado, E. N.; Lewis, R. A.; Faraone-Mennella, J.; Tezcan, F. A. *J. Am. Chem. Soc.* **2008**, *130*, 6082–6084.
(243) Kimura, E.; Shiota, T.; Koike, T.; Shiro, M.; Kodama, M. *J. Am. Chem. Soc.* **1990**, *112*, 5805–5811.
(244) Bazzicalupi, C.; Bencini, A.; Bianchi, A.; Fusi, V.; Giorgi, C.; Paoletti, P.; Valtancoli, B.; Zanchi, D. *Inorg. Chem.* **1997**, *36*, 2784–2790.
(245) Koerner, T. B.; Brown, R. S. *Can. J. Chem.* **2002**, *80*, 183–191.
(246) Koike, T.; Takamura, M.; Kimura, E. *J. Am. Chem. Soc.* **1994**, *116*, 8443–8449.
(247) Kimura, E.; Hashimoto, H.; Koike, T. *J. Am. Chem. Soc.* **1996**, *118*, 10963–10970.
(248) Innocenti, A.; Scozzafava, A.; Parkkila, S.; Puccetti, L.; De Simone, G.; Supuran, C. T. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2267–2271.
(249) Nomura, A.; Sugiura, Y. *J. Am. Chem. Soc.* **2004**, *126*, 15374–15375.
(250) Patel, K.; Srivastava, K. R.; Durani, S. *Bioorg. Med. Chem.* **2010**, *18*, 8270–8276.
(251) Kiefer, L. L.; Fierke, C. A. *Biochemistry* **1994**, *33*, 15233–15240.
(252) Kiefer, L. L.; Ippolito, J. A.; Fierke, C. A.; Christianson, D. W. *J. Am. Chem. Soc.* **1993**, *115*, 12581–12582.
(253) Alexander, R. S.; Kiefer, L. L.; Fierke, C. A.; Christianson, D. W. *Biochemistry* **1993**, *32*, 1510–1518.
(254) Ippolito, J. A.; Christianson, D. W. *Biochemistry* **1994**, *33*, 15241–15249.
(255) Ippolito, J. A.; Christianson, D. W. *Biochemistry* **1993**, *32*, 9901–9905.
(256) Ippolito, J. A.; Baird, T. T.; McGee, S. A.; Christianson, D. W.; Fierke, C. A. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 5017–5021.
(257) Krebs, J. F.; Fierke, C. A.; Alexander, R. S.; Christianson, D. W. *Biochemistry* **1991**, *30*, 9153–9160.

(258) Kiefer, L. L.; Paterno, S. A.; Fierke, C. A. *J. Am. Chem. Soc.* **1995**, *117*, 6831–6837.
(259) Lesburg, C. A.; Christianson, D. W. *J. Am. Chem. Soc.* **1995**, *117*, 6838–6844.
(260) Huang, C. C.; Lesburg, C. A.; Kiefer, L. L.; Fierke, C. A.; Christianson, D. W. *Biochemistry* **1996**, *35*, 3439–3446.
(261) Elleby, B.; Sjöblom, B.; Lindskog, S. *Eur. J. Biochem.* **1999**, *262*, 516–521.
(262) Höst, G.; Mårtensson, L.-G.; Jonsson, B.-H. *Biochim. Biophys. Acta* **2006**, *1764*, 1601–1606.
(263) Höst, G. E.; Jonsson, B.-H. *Biochim. Biophys. Acta* **2008**, *1784*, 811–815.
(264) The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.
(265) Iranzo, O.; Thulstrup, P. W.; Ryu, S.-B.; Hemmingsen, L.; Pecoraro, V. L. *Chem.--Eur. J.* **2007**, *13*, 9178–9190.
(266) Dieckmann, G. Use of metal-binding de novo-designed alpha-helical peptides in the study of metalloprotein structures, University of Michigan, 1995.
(267) Ghosh, D.; Lee, K.-H.; Demeler, B.; Pecoraro, V. L. *Biochemistry* **2005**, *44*, 10732–10740.
(268) Gomis-Rüth, F. X.; Kress, L. F.; Bode, W. *EMBO J.* **1993**, *12*, 4151–4157.
(269) Avvaru, B. S.; Kim, C. U.; Sippel, K. H.; Gruner, S. M.; Agbandje-McKenna, M.; Silverman, D. N.; McKenna, R. *Biochemistry* **2010**, *49*, 249–251.
(270) Hunt, J. A.; Fierke, C. A. *J. Biol. Chem.* **1997**, *272*, 20364–20372.
(271) Fierke, C. A.; Calderone, T. L.; Krebs, J. F. *Biochemistry* **1991**, *30*, 11054–11063.
(272) Krebs, J. F.; Ippolito, J. A.; Christianson, D. W.; Fierke, C. A. *J. Biol. Chem.* **1993**, *268*, 27458–27466.
(273) Liang, Z.; Xue, Y.; Behravan, G.; Jonsson, B. H.; Lindskog, S. *Eur. J. Biochem.* **1993**, *211*, 821–827.